

UNIT 1: STATISTICS

GRAPHICAL REPRESENTATIONS

Graphical Representations Lessons and Exercises

Lesson	Assignment	Date
Intro + 1. Types of data (pg.5) 4-		
2. Bar Graphs (pgs. 6-13)		
3. Histograms (pgs. ^{14-18, 22} 12-15)		
4. Frequency Tables (pgs. ¹⁹⁻²¹ 16-17)		
5. Line Graphs (pgs. ^{23-27, 30} 18-22)		
6. Interpolation and Extrapolation (pgs. 28-29)		
7. Circle Graphs (pgs. ³¹⁻³⁵ 25-27)		
8. Comparing Types of Graphs (pg. 28) 36-		
Intro to Graphs + 9. Misleading Data (pgs. ⁴¹ 29 -32) 45		

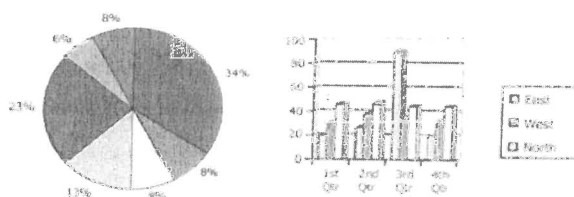
Introduction:

This unit focuses on displaying and interpreting data using various graphs. It also focuses on how to select the most suitable type of graph to present different types of data. Finally, this module shows how graphs are used to emphasize a certain point of view.

Sometimes a picture is easier to understand than a paragraph. Similarly, sometimes a graph is easier to understand than a list of numbers. The type of graph drawn usually depends on the kind of data one wants to illustrate.

The **4 types of graphs** that will be studied this unit are:

1. Bar Graphs
2. Histograms
3. Line Graphs
4. Circle Graphs



Choosing a scale and interval for bar, histogram, and line graphs (graphs with axes)

- For bar, histogram & line graphs, choose the scale of the axes carefully. Choose an appropriate scale so that all the numbers in the range will fit (it must *include all the values in the range*).
- The scale should be reasonable so that your graph is *as large as possible* for the space given and for your data.
- The interval must be uniform so that *each box on the axis must represent the same number across the whole axis* (ie label evenly, counting by 1, by 2, by 10, by 1000... for the entire axis - *don't change partway up the axis*). The two axes can have different scales from each other but the scale on one axis doesn't change. Also try to use interval that will make it the *easiest to graph the kind of numbers in the data*

How to determine Intervals

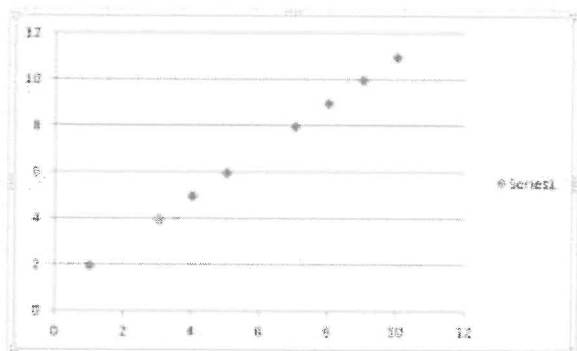
Favorite Singer	Number of Teachers
Toby Keith	22
Madonna	15
Elvis	11
Sting	5
Sinatra	2

- The interval is decided by your scale.
- In this case your scale would be from 2 – 22 and you want the scale to fit the graph.
- The best interval would be to go by 5's.

Lesson #1: Types of Data

Graphs represent either discrete or continuous data:

1. Discrete data is something you can count the whole number of.

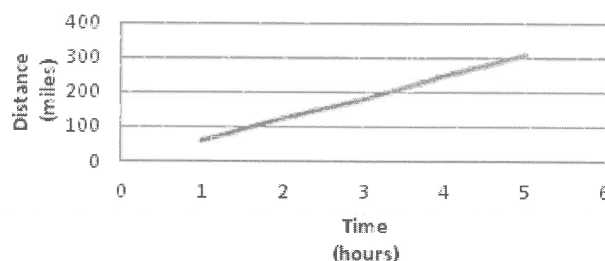


The units of measurement, such as # of CDs, cannot be split up; there is nothing between 1 CD and 2 CDs, nor can you purchase 2.5 CDs. There are no decimals or fractions that exist between the whole numbers. They are distinct and specific values, so the graph will only have points representing the data. We do not connect the points with a line.

2. Continuous data is something you can measure, such as time, temperature, distance, and weight.

The scale of measurement, such as distance, has meaning at all points between the numbers given. For example, we can travel a distance of 1.2 or even 1.63 kilometers. There are decimals or fractions between the whole numbers, so the points are connected with a line representing the relationship between the variables.

Distance vs. Time



Example: Place a 'D' for 'discrete' or 'C' for 'continuous' next to the types of data:

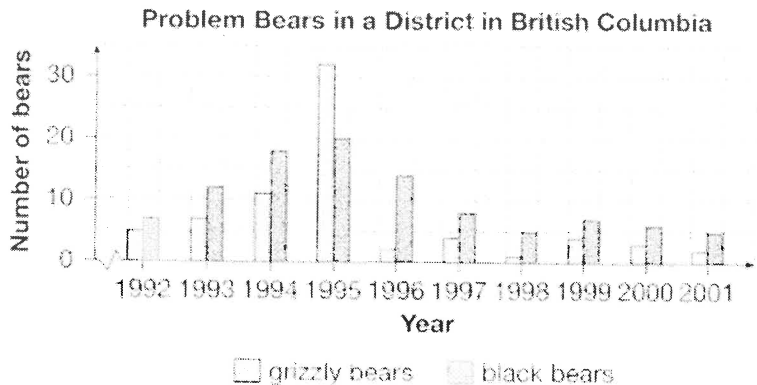
- a) D someone's height in centimetres
- b) D number of languages someone speaks
- c) C temperature range in a day *there are temps between measured temps*
- d) C speed of a train between cities *there is speed between measured speeds*
- e) D number of cars in the parking lot, sorted by color
- f) D favorite subject in school

Lesson #2: Bar Graphs

In this lesson, you will:

- Study bar graphs and the type of data they represent
- Create bar graphs from data
- Explore the advantages and disadvantages of using bar graphs to represent data

Amy is a conservation officer. She educates people about "problem bears." These are bears close to food in populated areas. What information can Amy get from this graph?



- 1 Describe the data in this double-bar graph.

The categories on the horizontal axis are

years from 1992-2001

The two groups of data are grizzly bears and black bears

trend

a relationship between two variables

- 2 Describe the **trend**. As the years increased, the number of problem black bears increased until 1995. Then it decreased, except for 1999.

range

the difference between the greatest and least number in a set of data

- 3 What is the **range** for each of these?

problem black bears: $20 - 5 = 15$

problem grizzly bears: $32 - 1 = 31$

- 4 Circle the correct word.

- The greater range means that the number of problem grizzly bears changed more ~~less~~ than that of black bears.

Bar graphs use vertical or horizontal bars to display discrete data sorted into well defined categories that are clearly separate from one another.

A bar graph is used for comparing quantities of many items. It is suitable for displaying data because each bar represents a quantity.

All bar graphs have both a vertical axis and a horizontal axis.



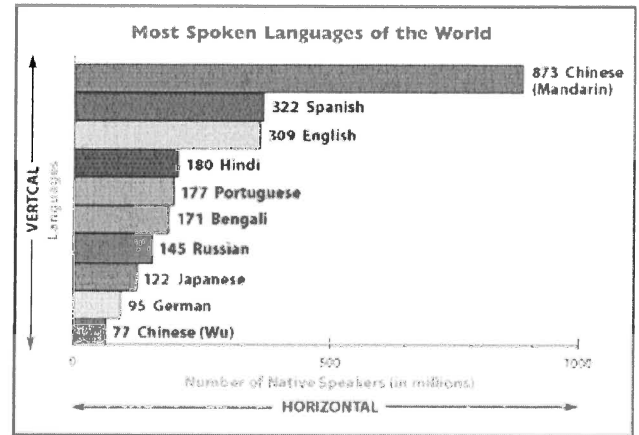
The above graph is a vertical bar graph because the bars are drawn vertically. The vertical or y-axis is the quantitative axis describing the amount. It is usually labelled, 'Number of...', because it refers to the number of items. The horizontal or x-axis is the qualitative axis, divided into categories.

NOTE: On a horizontal bar graph, the bars are drawn horizontally. In this case, the horizontal axis would be the quantitative axis, and the vertical axis would be the qualitative axis.

Drawing Bar Graphs:

Bar graphs must have the following:

- A title for the graph
- A title and *labels* for the qualitative axis
- A title and *scale* for the quantitative axis
- *Accurately* drawn bars with a space
- A *legend* (only if you are drawing a double bar graph)



To draw a good bar graph, you need to follow 5 steps:

1. Be sure the data is discrete and sorted into categories.

Question: What is your favorite subject?

Phys Ed: 179 students

Math: 14 students

Art: 28 students

French: 51 students

Music: 106 students

English: 65 students

Science: 47 students

2. Label the axes. Determine the qualitative and quantitative axes and label them.
3. Determine the maximum value of the sorted data. By knowing the highest value of any category, you know how long the longest bar will be. The quantitative axis should always start at zero, and should extend slightly past the maximum value.

Highest value is phys ed: 179 students

4. Determine the scale of the quantitative axis. You need to label a reasonable number of equal increments (not too few or too many). *Use as much of given space as possible*
The maximum value is close to 180. Increments of 10 or 20 would be reasonable.

5. Draw a bar for each category accurately. Each bar should be the same width, and there should be the same distance between each bar.

- Don't forget the title!
- Draw carefully. Use a ruler to draw the top lines (and side lines if not using graph paper) and to shade in.

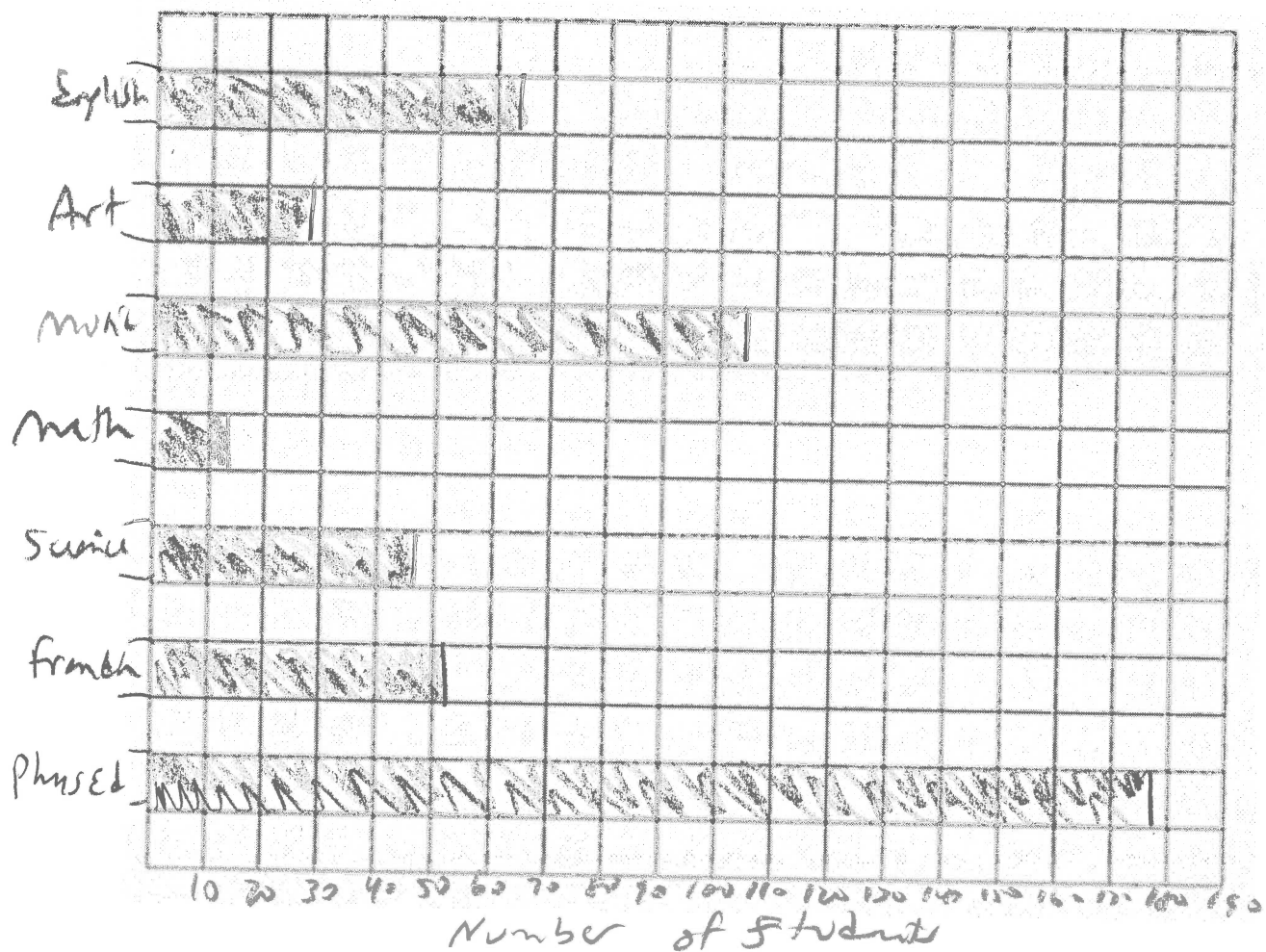
Try it! Create a vertical bar graph to represent the data on page 7.

Student Subject Preference



Try it! Create a horizontal bar graph to represent the same data.

Student Subject Preference

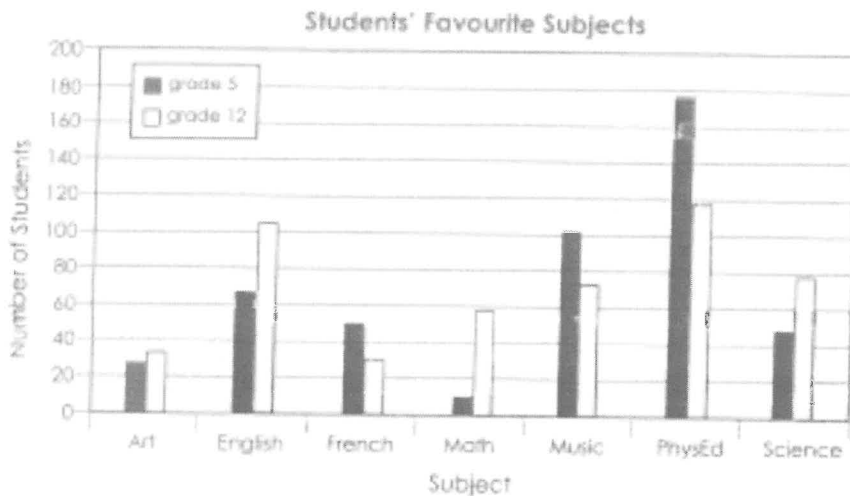


Double Bar Graphs:

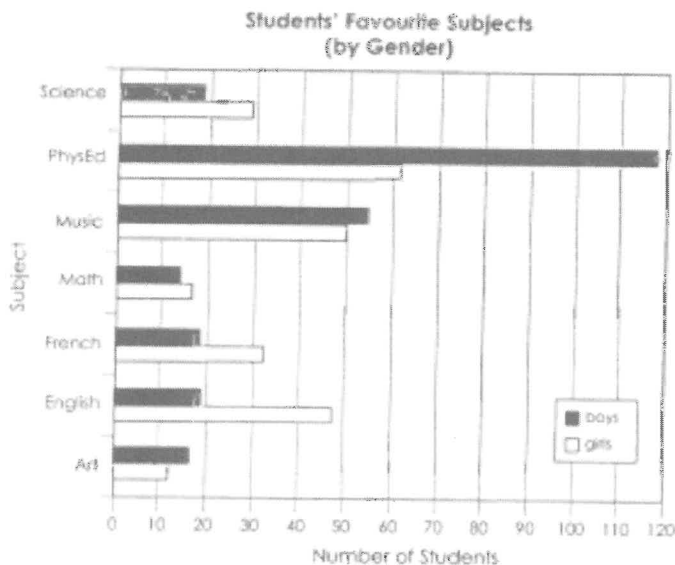
Double bar graphs compare data from two different surveys or different years, or highlight a difference that is in every category, such as boys vs girls.

NOTE: The steps for drawing a double bar graph are the same as drawing a bar graph, however, a **legend must** be included to distinguish between the two different data sets.

Go back to the example of students' favorite subjects. The survey could be conducted on Grade 5 students and on Grade 12 students, and then compared.



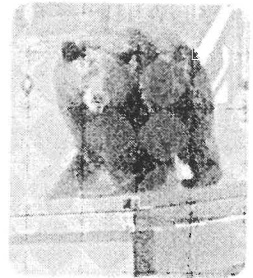
Another possibility would be to break down the data from the Grade 5 students into boys and girls.



Try it! Create a vertical double bar graph to represent the data that follows.

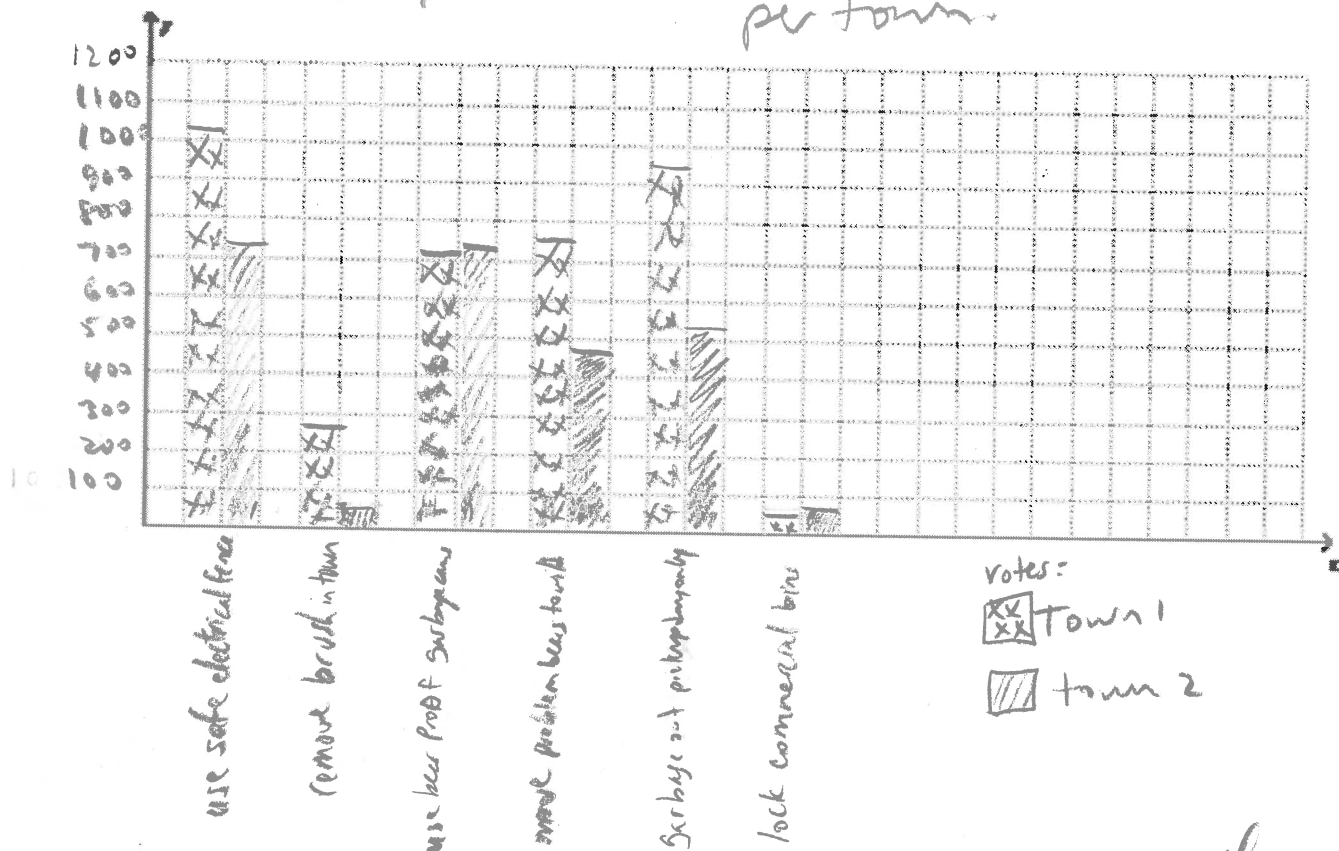
Example

Councils in two British Columbia towns conducted a survey. They wanted to choose three ways to protect the bears and keep communities safe. The results are shown in the chart. What decision might they make from the data?



Suggestion	Votes: Town 1	Votes: Town 2
use safe electric fence around landfill	1020	711
remove brush in town	294	47
use bear proof garbage cans	701	710
move problem bears to the wild	773	479
put out garbage on pickup day only	948	518
lock commercial garbage bins	60	76

Safe communities + bear survey results
per town



decision - fence because highest # of votes from both towns.

Some Advantages and Disadvantages of Bar Graphs

advantages	disadvantages
<ul style="list-style-type: none"> • Eye catching • Easy to understand • Easy to compare sets of data or facts (in double and triple bar graphs) quickly • The bars provide a visual display for comparing quantities in different categories or groups.. 	<ul style="list-style-type: none"> • Only for discrete data • Shows the number of items in a category (not a percentage) • You have to estimate the length of the bar so that it shows the correct value (For example if your axis is in intervals of 1000, how do you know where exactly to draw the bar showing 832?) • Can be difficult to read accurately • If range of data is large, some bars may be very share and hard to estimate or even hard to see. (If your axis is in intervals of 1000, how teeny is the draw a bar showing 5? Can you even see it?) • Can be misleading if scale does not begin at zero. <div data-bbox="500 951 1308 1455"> <p>correct version vs. incorrect</p> </div> <ul style="list-style-type: none"> • Can be misleading if sample sizes are not equal. <div data-bbox="592 1549 1409 1837"> <p>Same data.</p> </div>

Lesson #3: Histograms

In this lesson, you will:

- Study histograms and the type of data they represent
- Create histograms from data
- Explore the advantages and disadvantages of using histograms to represent data

frequency table

a table that shows the number of items in each interval

Hint

Each interval includes numbers that are greater than the lesser value and up to, and including, the greater value.

Matt works at the post office. This frequency table shows the masses of parcels he weighed for postage last month. What does this histogram tell you about the masses?

Mass (g) (over-including)	Number of parcels
500-1000	292
1000-1500	282
1500-2000	287
2000-2500	233
2500-3000	214
3000-3500	236



histogram

a graph that shows data organized into intervals of equal size; the touching bars show the frequency

- Describe the data in this histogram.

The horizontal axis shows masses of parcels. The width of each interval represents 500 g. The vertical axis shows the frequency (number of parcels)

- Circle the correct word or phrase to complete each sentence.

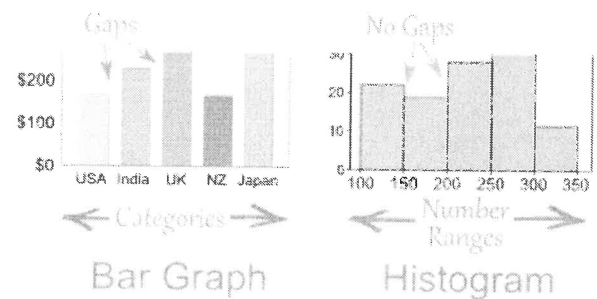
For each of the first three intervals, the number of parcels is about the same / a wide range. For each of the last three intervals, the number of parcels is more / less than 250.

- Can Matt tell the exact mass of any parcel from the histogram? no

What does the histogram tell Matt about mass? There are about 250 parcels in each 500 g interval. The least possible mass is just over 500 g. The greatest possible mass is 3500 g.

The Parts of a Histogram:

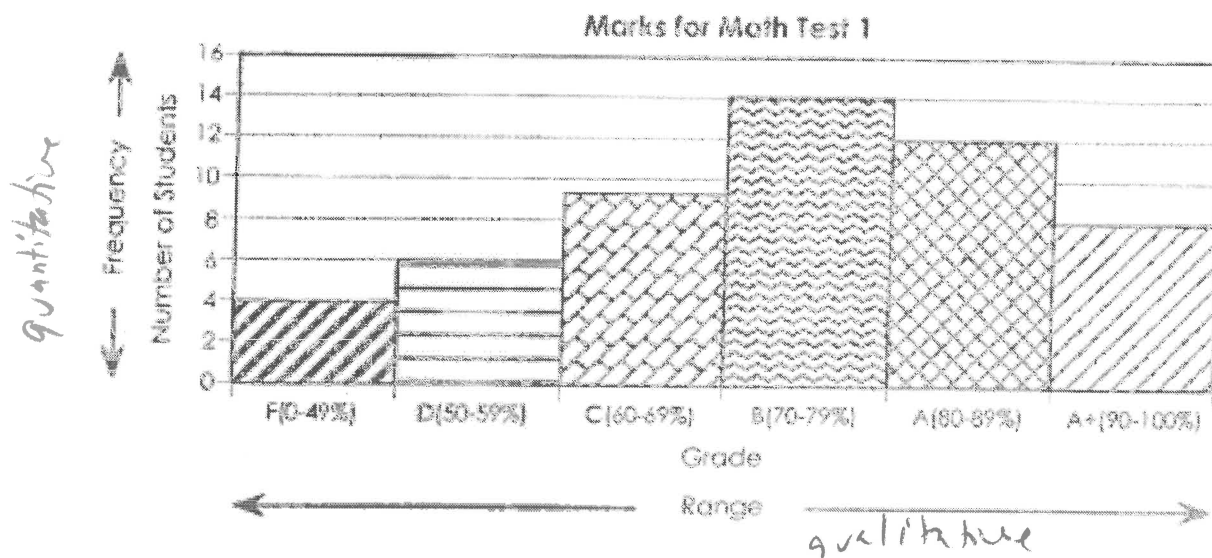
A **histogram** is a bar graph which shows **frequency distribution**. While histograms look a lot like bar graphs, and are created in much the same way, there are some important differences between the two.



Bar graphs display discrete data (example one bar for each student). Histograms use vertical or horizontal bars to display continuous data (data values that are connected) that is sorted into categories that represent a range of data. Continuous data means that each category represents a range of values, and it is possible to have any value within the range.

*Note: The bars do not have any gaps between them. The bars touch because the data is continuous. Note this on your resource sheet.

All histograms have both a vertical axis and a horizontal axis.



The above graph is a vertical histogram because the bars are drawn vertically. The vertical or y-axis is the quantitative axis which shows the frequency. The horizontal or x-axis is the qualitative axis that displays the range.

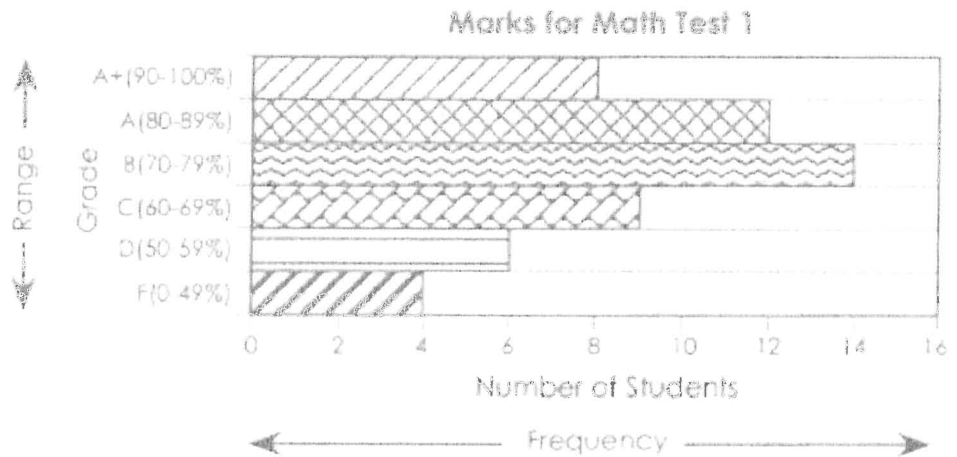
→ The range of each category is the difference between the minimum and maximum values.

NOTE: Similar to a bar graph, on a horizontal histogram, the axes are switched. The horizontal axis is the quantitative axis, and the vertical axis is the qualitative axis.

Drawing Histograms:

Histograms must have the following:

- A title for the graph
- A title and **labels (ranges)** for the qualitative axis
- A title and **uniform scale** for the quantitative axis (*the frequency*)
- **Accurately** drawn bars, with **NO** spaces between them



To draw a good histogram, you need to follow 4 steps:

(TAILS 1-7)

1. Sort the continuous data into sections so that all possible data is included.
In the above graph, the sections range from 0 to 100%. No one necessarily received 0% or 100%, but these are the minimum and maximum marks available.

You do not need to divide the sections evenly across the range. As you can see, the category for an F is 0-49%, while the other categories have much smaller ranges. This is done to avoid having very small bars which can become hard to read and do not communicate data very well.
 2. Determine the maximum frequency of the sorted data. By knowing the highest value of any category, you know how long the longest bar will be.
 3. Draw and label the vertical and horizontal axes. Always draw the axis describing frequency to extend further than the maximum value.
 4. Draw a bar for each category **accurately**. Each bar should be the same width, and there should be **NO** spaces between bars.
- Don't forget the title!
 - Use a ruler!

Try it!

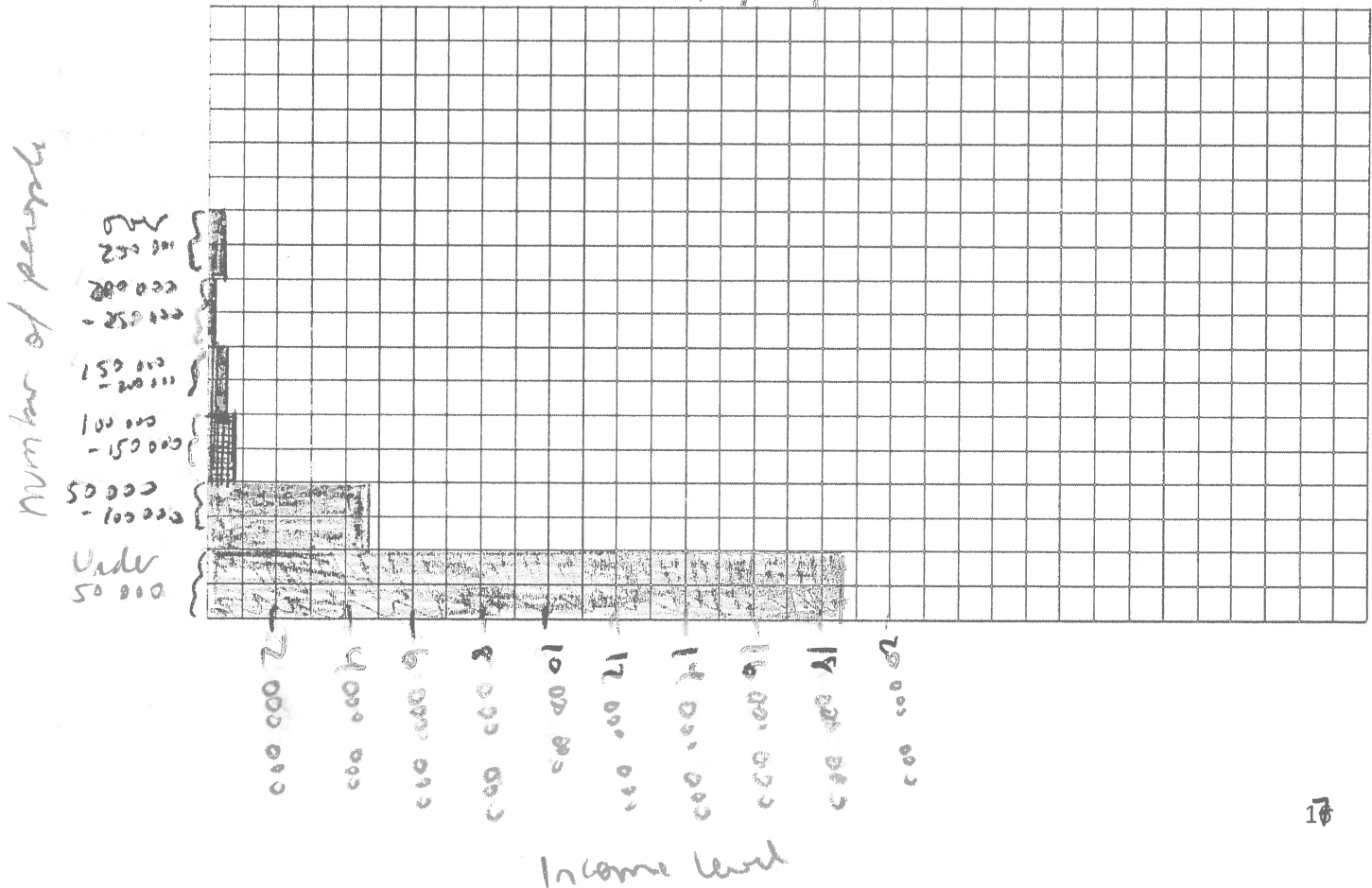
In 2007, Statistics Canada did a survey to determine the incomes of employed people in Canada. The results, after compiling the information, are shown below.

Ranges of Income	Number of People in Each Range
under \$50,000	18 395 690
\$50,000-\$100,000	4 759 410
\$100,000-\$150,000	756 930
\$150,000-\$200,000	197 130
\$200,000-\$250,000	82 020
over \$250,000	160 060

← max frequency.

Create a horizontal histogram to represent this information.

number of people per income level



Try it!

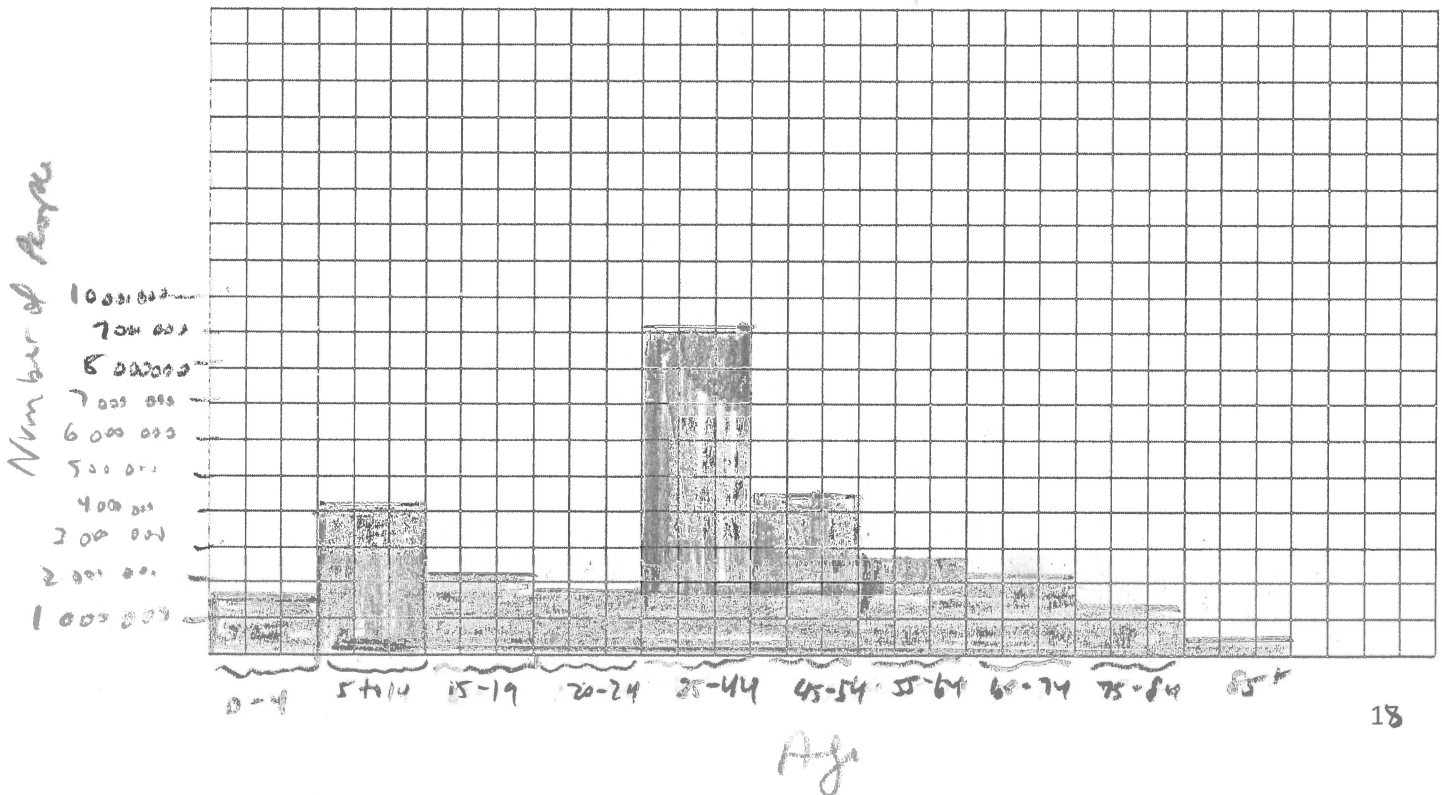
In 2001 Statistics Canada surveyed the population to find out how many people lived in Canada. The results were as follows:

Create a vertical histogram to represent this data.

Age	Number of People
0 to 4	1,696,285
5 to 14	4,029,255
15 to 19	2,053,325
20 to 24	1,955,810
25 to 44	9,096,560
45 to 54	4,419,290
55 to 64	2,868,015
65 to 74	2,142,835
75 to 84	1,329,810
85+	415,910

What conclusions can you make based on the graph? the most common ages are 25-44.
 not the majority

Population of Canada



Creating Histograms - Construct Frequency Table

Often the histograms you will draw will have the first 3 steps completed and the frequency chart drawn. But if you are presented with raw data, you need to create a tally chart and then a frequency chart.

1. Gather the data you are interested in.

The example data here represents the heights of students in Grade 11 in centimetres.

152 161 189 158 177 167 182 155 171 163 185 173 183
 166 177 172 157 168 181 167 188 158 168 162 159 153
 171 186 152 167 173 175 159 184 189 187 162 151 179
 163 185 183 177 176 152 174 179 182 158 172 186 171
 185 157 173 188 152 176 187 157 172 179

50 + 12

62

2. Look at the data to determine the categories or intervals you will use to organize the data. (Sometimes the categories/intervals will be given to you but otherwise you must determine them from the data.

- i. Find the range (largest value – smallest value): $189 - 151 = 38$.
- ii. How many groups do you want? *example: 4 groups*
- iii. 3. Divide your range by the number of groups you want.

$$38 \div 4 = 9.5. \text{ Round up to } 10$$

- iv. 4. Pick a start value. *152 is the smallest number so start with 150.*
- v. Try it. Make a tally chart with your intervals. See if your last value works out with 4 groups with intervals of 10 cm. If you don't like the groups, go back and change the group size or the starting value.
- vi. Make a tally chart to count the number of people with heights in each category. (Every 5th line is diagonal through the 4 lines. That way in the end you can count by 5 to find the total.) *make sure the data = the total frequency.*

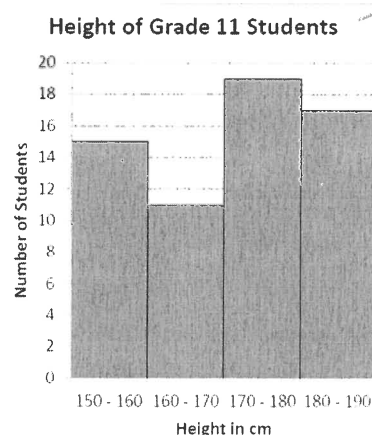
Height Classes	Tally
150 - 160 cm	
160 - 170 cm	
170 - 180 cm	
180 - 190 cm	

Height (cm)	Number of Students
150 - 160	15
160 - 170	11
170 - 180	19
180 - 190	17

3. Using the totals from your tally chart, construct a frequency table for your data. The frequency corresponds to the number of times each value is observed. THE REST OF THE STEPS ARE HOW TO CREATE A HISTOGRAM. IF A FREQUENCY TABLE IS PROVIDED FOR YOU, START WITH THESE STEPS.

4. To create a histogram, draw a horizontal and a vertical axis. Decide which axis is quantitative (frequency – *example number of students*) and which axis is qualitative (categories or intervals – *example heights*). Label your axes with numbers/ categories and titles.

5. For each category/interval, draw a rectangle (without space between the rectangles). The width of the rectangle represents the interval between two groups, and the height represents the observed frequency.



Frequency Tables:**Example**

This data shows the number of acres on 32 sugar beet farms near Taber. Quinn's family wants to grow sugar beets on their farm near Taber.

How can Quinn's family use this data to help them decide how many acres they should use for sugar beets?

139	61	358	169
126	350	62	159
502	290	150	74
61	462	59	122
187	72	76	66
123	66	150	191
130	145	150	231
398	836	208	420

Solution

- A. Organize the data into intervals to create a frequency table. Choose the number and width of the intervals.

- Determine the range. $836 - 59 = 777$
- Choose an appropriate number of intervals.
Suppose you use 100 for the width of each interval.
Divide 777 by 100 and round up to get 8 intervals.
- Start with the first interval. Choose a value that is lower than the minimum number. The minimum is 59.

$$777 \div 100 = 7.77$$

- B. Complete the first row in the frequency table using intervals of 100. Then complete the second row.

Acres of sugar beets (over-including)	50-150	150-250	250-350	350-450	450-550	550-650	650-750	750-850
Frequency (number of farms)	18	6	1	4	1	1	0	1

less than 59

|||||

|||||

|||||

|||||

more than 836

32

Try it!

Jill is a tour guide in Yellowknife. She takes tour groups to nearby Cameron Falls in July. She needs to let the tourists know what temperatures to expect.

- a) Use the data below to create a frequency table.

Temperatures from Previous Year (°C)			
13.8	13.5	15.7	16.7
13.9	16.0	13.4	18.4
15.2	18.3	16.2	19.4
13.4	20.9	17.3	19.4
12.6	23.2	19.4	21.6
12.6	21.3	21.3	20.4
16.9	20.4	21.2	20.4
19.1	17.8	20.9	

24
331

Average Daily Temperatures in Yellowknife in July	
Temperature (°C) (over-including)	Frequency (number of days)
10-13	2
13-16	7
16-19	8
19-22	13
22-25	1
	<u>31</u>

$$23.2 - 12.6 = 10.6$$

$$\frac{10.6}{5} = 2.12$$

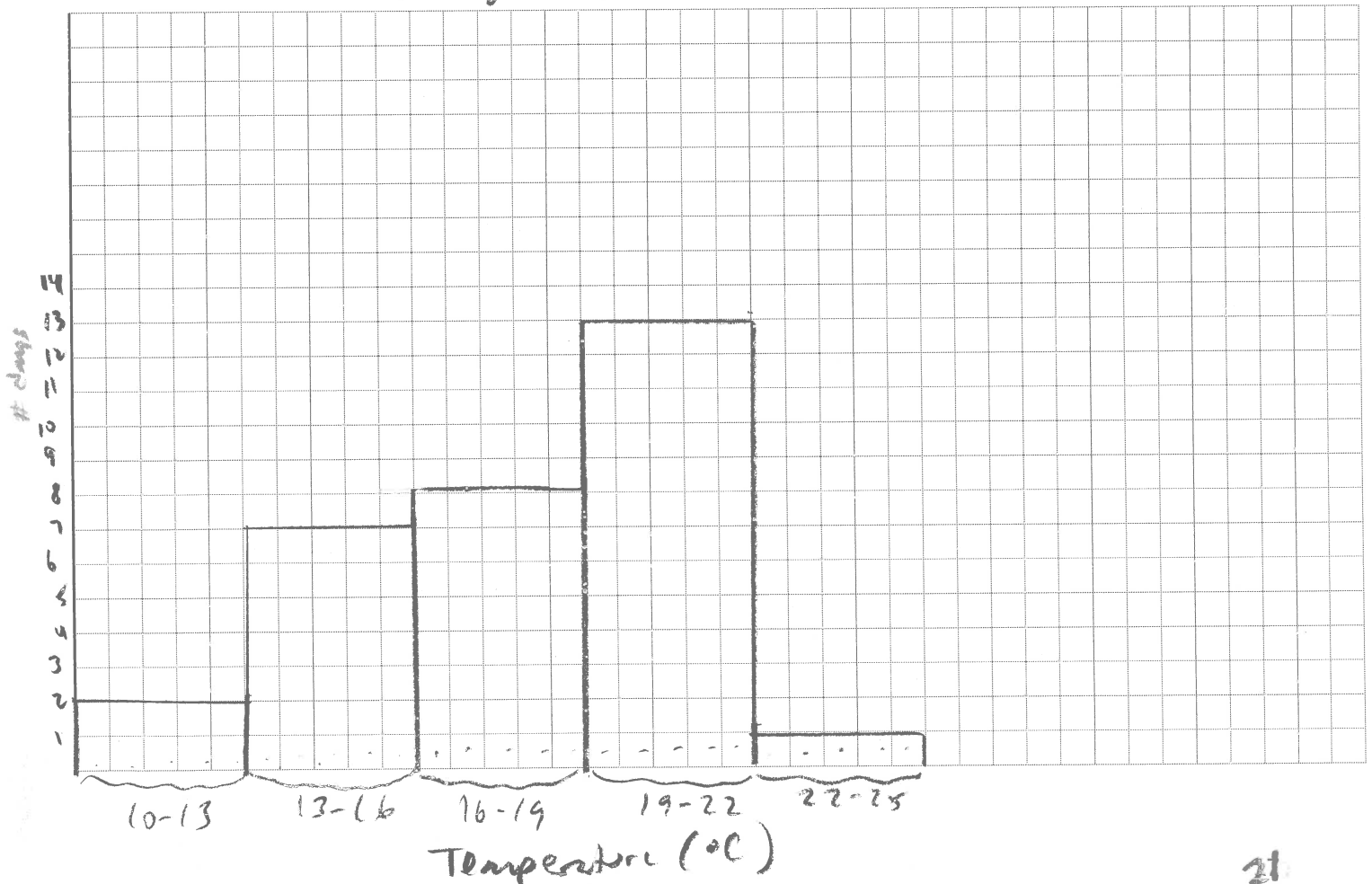
Hint

Each temperature interval must be the same size.

11
+++
+++
+++

- b) Create a histogram on grid paper.

Average Daily Temperature in Yellowknife in July.



Some Advantages and Disadvantages of Histograms

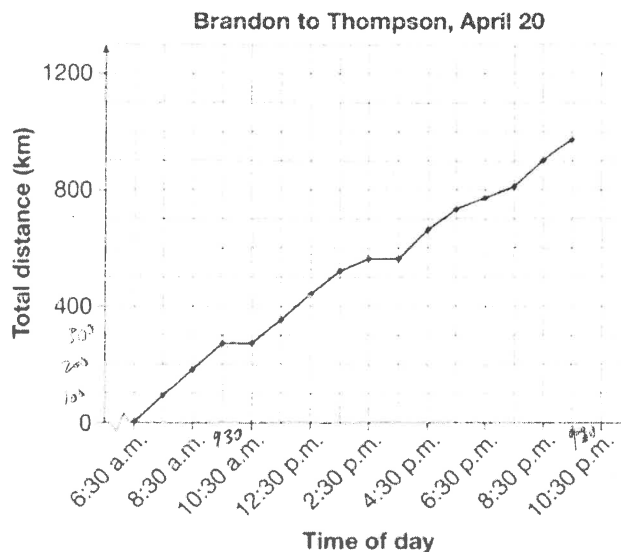
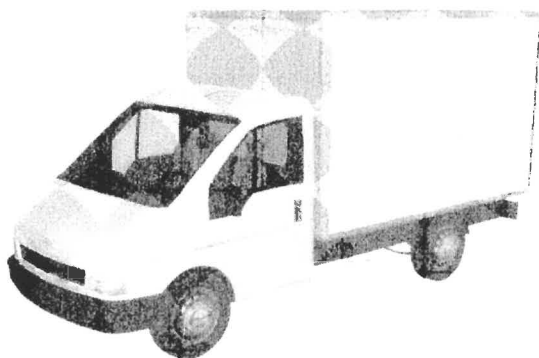
<u>advantages</u>	<u>disadvantages</u>
<ul style="list-style-type: none"> • Summarize data • Group data into categories so that it is easier to see patterns • Useful when describing a range of data all related to one topic (<i>ex. Income of Canadians</i>) 	<ul style="list-style-type: none"> • Only for use with continuous data • Do not include exact data values (<i>ex from the graph in example on page 19, you cannot tell how many people are 25 or how many people are 44</i>) • Hard to compare with other data if both graphs do not have the same categories/sections, because you do not know how many data points from one category would be in another if you changed the range. (<i>ex from that same graph on p. 19, you can't divide "age 25-44" into two parts because you wouldn't know that (perhaps) only 15% of those people are older than 34</i>) • Can be misread. You must use correct words. (<i>ex. In that same graph on p. 19, some might erroneously think that the "majority" of people are between the ages of 25-44. Majority means more than 50% of the total. That age group is simply the most common. Notice too that 25-44 has an age range of almost 20 years while the others have 5 or 10 years. If you split it up, the two bars would be shorter. But we can't split it up because that's how the data was given - we don't have the raw data.</i>)

Lesson #4: Line Graphs

In this lesson, you will:

- Study line graphs and the type of data they represent
- Create line graphs from data
- Explore the advantages and disadvantages of using line graphs to represent data
- Interpolate and extrapolate values from a graph

Camden and Maria drove a delivery truck from Brandon to Thompson in a day. They created a graph from their notes about the trip. Describe their drive.



REFLECTING

Why does the line never show a decrease?

- 1 Describe the trend.

As time increases, the distance travelled either increases or stays the same

- 2 Does the graph show they stopped at 9:30 a.m.? yes

Explain. The line is horizontal (no km)

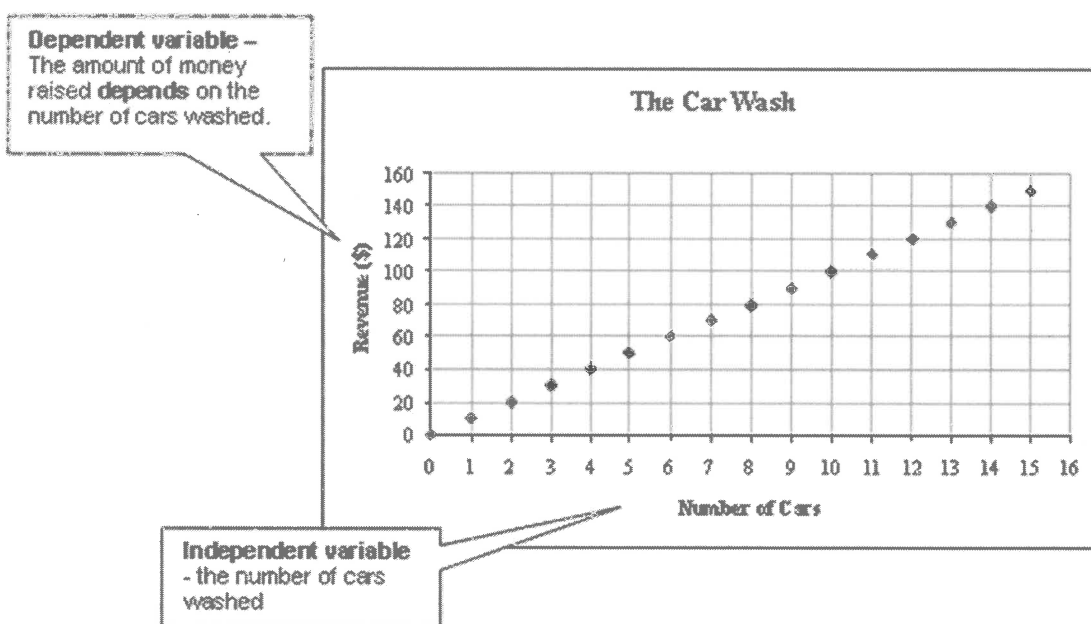
- 3 How far had they driven before 9:30 a.m.? about 280 km

- 4 How long did the trip take? about 15h

Axes: Independent and Dependent Variables

Line graphs, like histograms, display continuous data. A line graph provides a quick visual impression of the relation between two quantities or groups of data, such as the relationship between temperature and time. One axis describes the time of day, while the other describes temperature. These two characteristics are called variables.

- The independent variable is on the *horizontal or x-axis*. The independent variable stands alone and isn't changed or affected by the other variable you are trying to measure. For example, how many hours someone studied for a test might be an independent variable.



- The dependent variable depends on other factors. For example, a person's test score could be a dependent variable because it could change depending on how much time the person spent studying. Usually when you are looking for a relationship between two things, you are trying to find out what makes the dependent variable change the way it does.

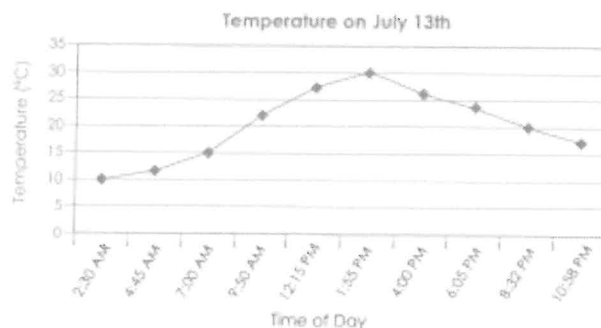
Example: Identify the dependent (D) and independent (I) variables in each statement:

- The distance you are from ^Ihome compared to the amount of time you spent driving ^Das you go to work
- The number of pizzas needed compared to the number of people at the party ^D^I
- The age compared to the height of a person ^I^D

Drawing a Line Graph:

Line graphs must have the following:

- A title for the graph
- A title on each axes
- **Scales** or **labels** for each axes
- **Accurately** plotted points, **connected** by a straight line



To draw a good line graph, you need to follow 5 steps:

(TAILS P-9)

1. Identify which variable is dependent and which is independent

*A graph shows the thickness of the ice on Lake Winnipeg each month of the year.
The thickness of the ice is the dependent variable, while the month is independent.*

2. Determine the range of data for both variables

October:	0 cm	November:	10 cm	December:	30 cm
January:	75 cm	February:	75 cm	March:	55 cm
April:	28 cm	May:	6 cm	June:	0 cm

The range from October to June is 0 - 75 cm.

3. Draw and label the axes. Make sure to use a **uniform** scale.
4. Plot the points on the graph. Make sure the each point is lined up with the correct values on the horizontal and vertical axes.

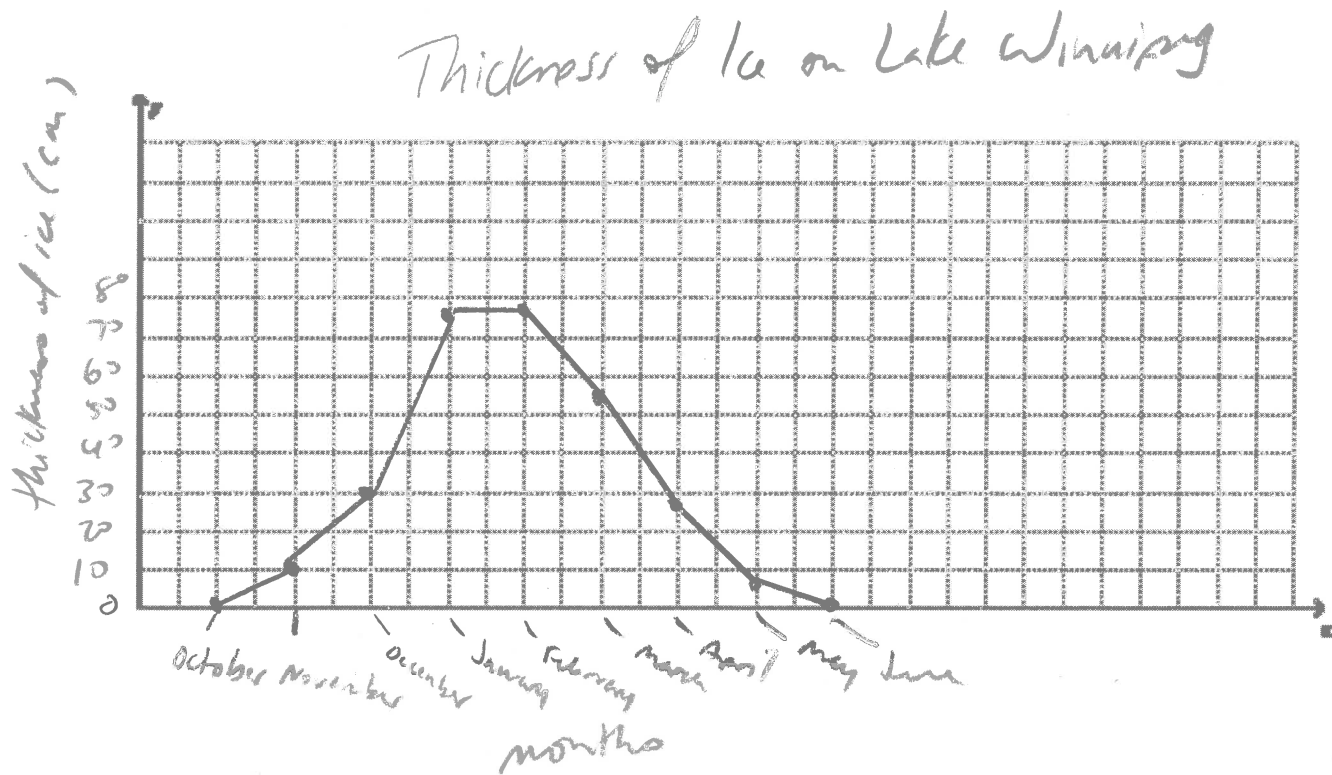
The first ordered pair is: (October, 0)

5. Connect the points using line segments

6. Don't forget the title!

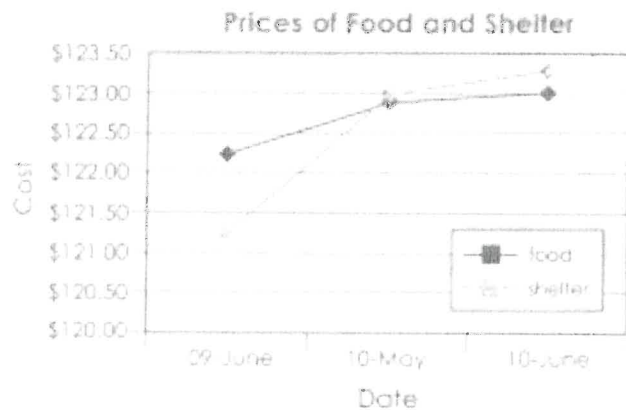
7. Rule

Try it! Create a line graph to represent the data on page 20.



Double Line Graphs:

Double line graphs are similar to double bar graphs, in that they are useful when making comparisons. The steps you follow for drawing a double line graph are the same as for drawing a line graph, however you need to determine the ranges of both sets of data before you start. You also need to include a legend.



Try it! Create a double line graph to represent the data below.

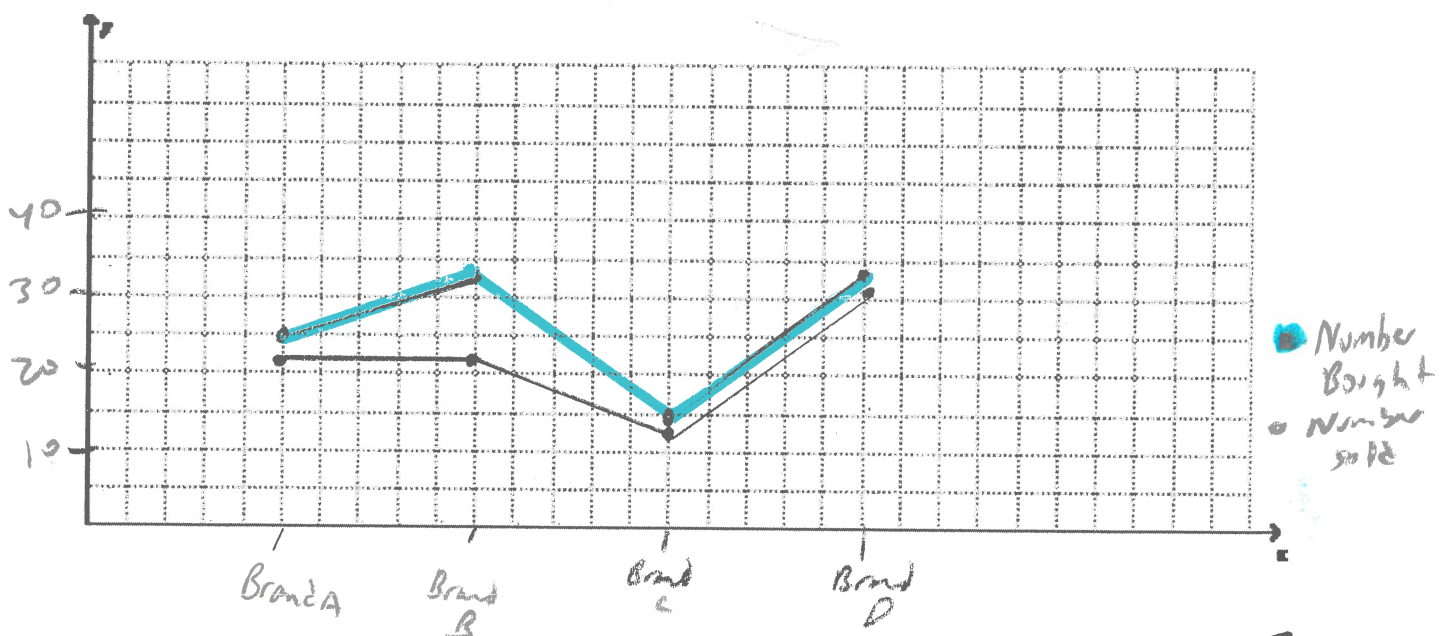
Tomas owns a boat shop and is tracking the sales of different brands of outboard motors. The table below shows the number of motors Tomas has ordered and the number of sales.

Type	Brand A	Brand B	Brand C	Brand D
Number bought	25	32	15	34
Number sold	21	22	14	32

range:

brand A-D 25-34

brand A-D 21-32



Try it!**Example 1**

Layla's truck can hold up to 1000 lb. She used an online converter to get these equivalent masses. How many kilograms of construction materials can she deliver in her truck?

Metric and Imperial Mass	
10 kg	22.05 lb
20 kg	44.09 lb
30 kg	66.14 lb
40 kg	88.18 lb
50 kg	110.23 lb

Solution

- A. Label the scale for mass in kilograms and the scale for mass in pounds. Plot the points in Layla's chart. Complete the line graph.

- B. What trend does the graph show?

As the mass in kilograms increases, the mass in pounds increases. The points lie in a straight line going up to the right.

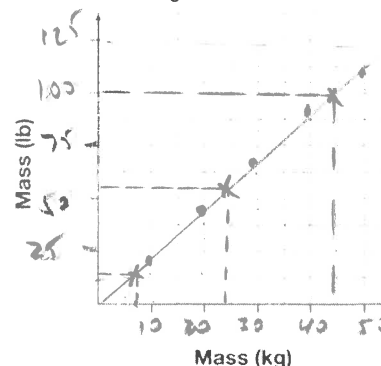
- C. **Interpolate** to convert these masses.

- a 7 kg box of nails: about 15 lb
- a 54 lb bag of cement: about 18 kg

- D. What is the total mass in kilograms that Layla's truck can carry?

On the graph, 100 lb is about 45 kg. So if you multiply 100 by 10 you know that 1000 lb is about 450 kg.

Kilograms and Pounds

**Example 2**

Shawna created this graph about the fuel economy of her hybrid vehicle. How far can she drive on a full tank of gas?

Solution

- A. How much gas did Shawna start with? 62 L of gas.

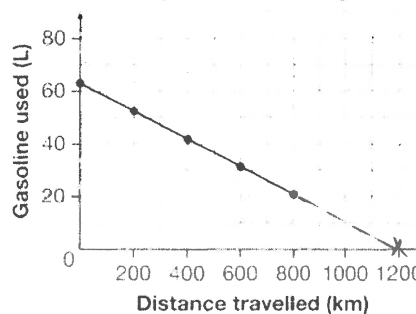
- B. What trend does the graph show?

As the distance travelled increases, the volume of gas decreases. The points lie in a straight line going down to the right.

- C. **Extrapolate** by extending the graph to show when the tank is empty.

Shawna does not buy any gas. Then she can drive about 1200 km before the tank is empty.

Hybrid Vehicle Range



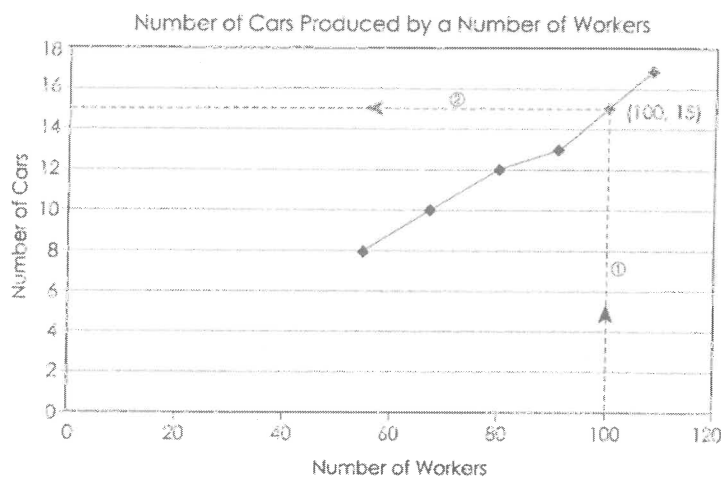
extrapolate
estimate beyond
known points

Interpolation and Extrapolation:

Unlike other graphs, line graphs provide information *in addition* to the data collected. You can use a line graph to predict points both inside and outside the range of data.

Interpolation means interpreting information from inside a set of data. It is estimating values that are *not* included in the data.

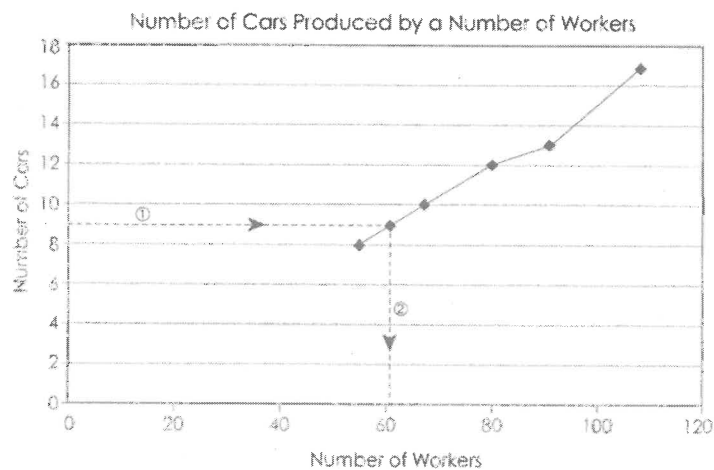
You could interpolate to determine the approximate number of cars produced with 100 workers. Draw a vertical line up from the 100 on the horizontal axis until it crosses the graph. Then draw a horizontal line from that point to the vertical axis.



Extrapolation is estimating information outside the range of data. You can estimate values that are *not* included in the data, and that lie outside of the graph by extending the lines.

You could extrapolate the approximate number of cars produced with 40 workers. Extend the line to the left of 55 people until it passes 40 people. Then follow the same steps you used with interpolation.

Try it!

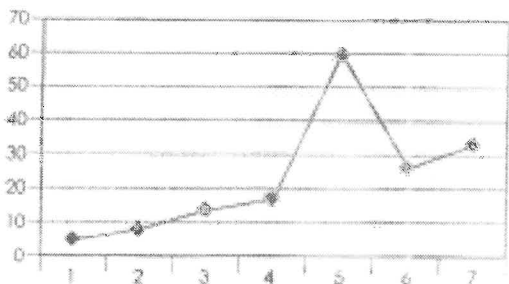


Some Advantages and Disadvantages of Line Graphs

advantages

- Easy to quickly analyse data that changes continuously over time
- Easy to compare two sets of data (double line graphs) or to see changes over time
- Makes relationships and overall trends within data clear (increasing, decreasing, or fluctuating - up and down)
- Shows range, minimum, maximum, and outliers clearly
- Can use with discrete or continuous data
- Useful for making predictions (interpolating or extrapolating)
- When data is displayed in line graph, any data that does not fit the trends becomes obvious.

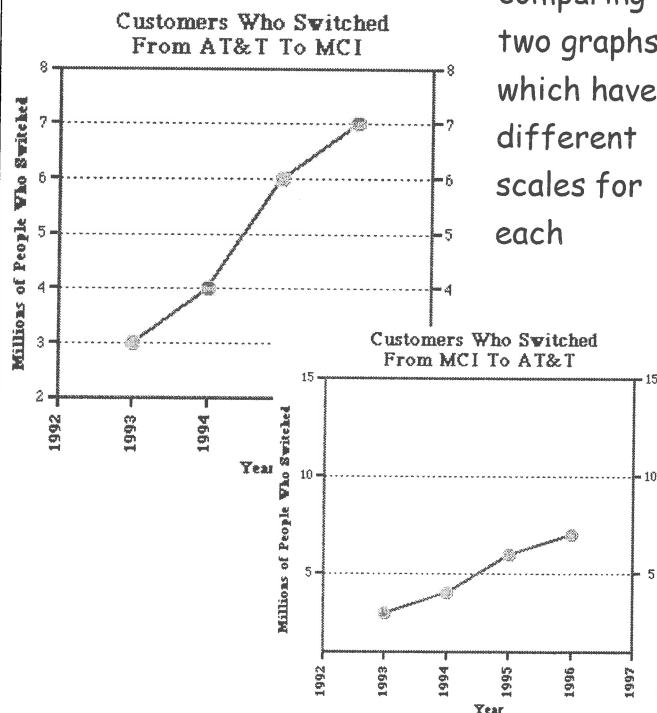
(Example in the line below, it's obvious that the point above 5 on horizontal axis does not fit in with the rest of the data.)



disadvantages

- If the trend of the data is not a straight line, then your estimation based on extrapolation could be less accurate. (Extrapolation can be less reliable than interpolation.)
- If there are too many fluctuations in the line, it can make the trend difficult to see.
- Needs small range of data
- Best for fewer than 50 data values
- Easy to alter or manipulate the way a line graph appears to make the data look a certain way and give a false impression (either by not using consistent scales on the axes or when

comparing two graphs which have different scales for each



Lesson #5: Circle Graphs

In this lesson, you will:

- Study circle graphs and the type of data they represent
- Create circle graphs from data
- Explore the advantages and disadvantages of using circle graphs to represent data

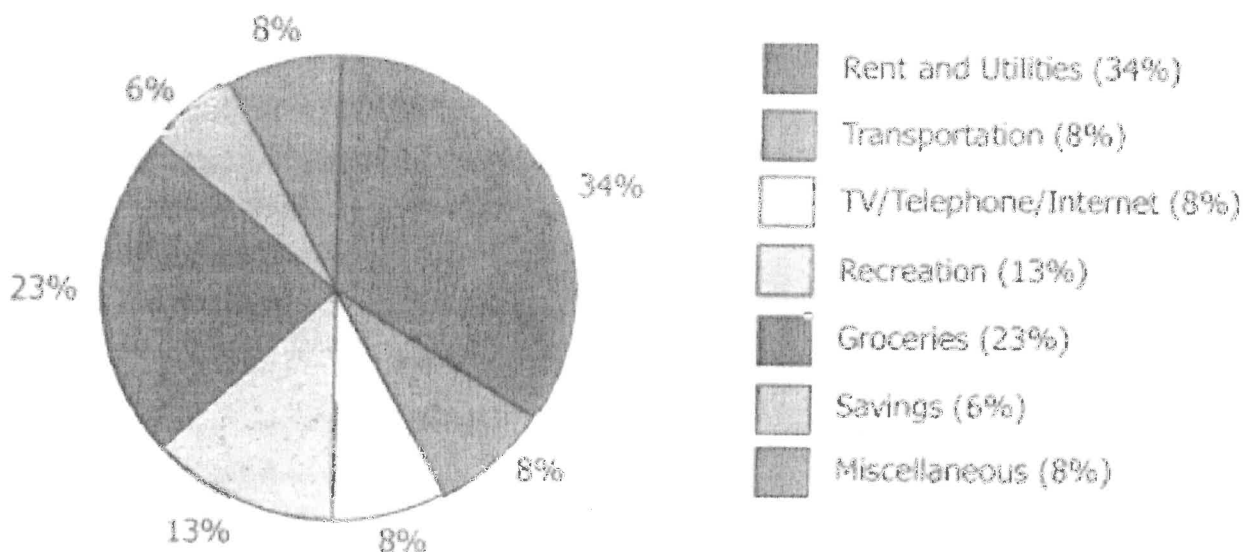
A **circle graph** represents well defined categories of **discrete data** as sectors of a circle (*like slices of a pizza*). All circle graphs compare **parts of a whole**, where the *whole* represents *100% of the data*. The sections of a circle graph are labelled with a **percentage**, which reflects how much of the data is represented by that sector.

Drawing Circle Graphs:

Circle graphs must have the following:

- A title for the graph
- **Accurately** drawn **sectors** (correct-sized angles drawn using protractor)
- **Percentage labels** on each sector (beside if they don't fit)
- A **legend** explaining what each sector represents, and the **actual values**
- **Accurate** division of the circle based on the percentage of each sector compared to the total population.

How the Paycheck Is Spent



To draw a good circle graph, you need to follow 4 steps:

1. Be sure the data is **discrete** (the same information cannot be placed into two or more categories)

Where are you in a 24 hour period, and how long are you there for? (The survey produces discrete data for the circle graph because you cannot be in two places at once.)

2. **Rewrite** each category of data as a percentage of the total. For each sector, take the actual value, divide by the total, and multiply by 100.

3 hours in the car, 8 hours at work, 2 hours at the gym, 11 hours at home

Total number of hours in a day is 24:

% in the car: $3/24 \times 100 = 12.5\%$

% at work: $8/24 \times 100 = 33.3\%$

% at the gym: $2/24 \times 100 = 8.3\%$

% at home: $11/24 \times 100 = 45.8\%$

NOTE: the sum of the percentages should be **approximately** 100% (may not be exact due to rounding)

$12.5 + 33.3 + 8.3 + 45.8 = 99.9\%$

3. **Convert** the percentage of each sector into degrees (360°). For each sector, convert the percentage back into a decimal (divide by 100), and multiply by 360° .

Car: $12.5\%/100 = 0.125 \times 360^\circ = 45^\circ$

Work: $33.3\%/100 = 0.333 \times 360^\circ = 120^\circ$

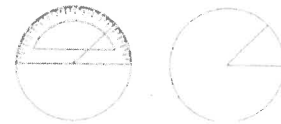
Gym: $8.3\%/100 = 0.083 \times 360^\circ = 30^\circ$

Home: $45.8\%/100 = 0.458 \times 360^\circ = 165^\circ$

4. Draw a circle with a compass, then use a **protractor** to draw each sector.



- Use a compass to accurately draw a circle
- Draw a radius from the center of the circle to the outside of the circle
- Measure the angle of the first sector from the radius using a protractor (ex. car 45°)
- Continue measuring the angles of the sectors from the previous sector
- Label each sector with their percent values (not degrees)
- Place the legend beside the circle, including percents and actual values
- Give your circle graph a title



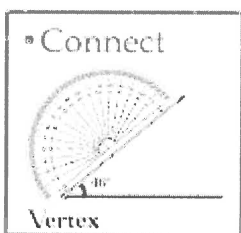
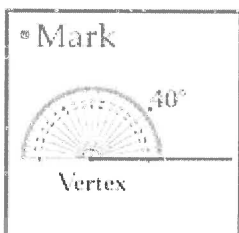
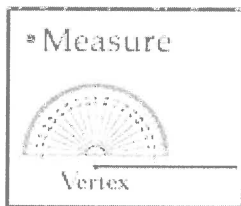
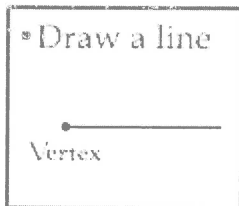
Include these 4 steps in your own words in your resource sheet.

- **Try it!** Create a circle graph to represent the data on page 32. Use a compass to draw the angles. Use the degrees in step 3 to draw the angles for the sectors.

Hint: you may find it helpful to turn the paper. Each time you draw an angle, that line becomes the starting point for your next angle.

Hint: measure the final angle to make sure it's accurate.

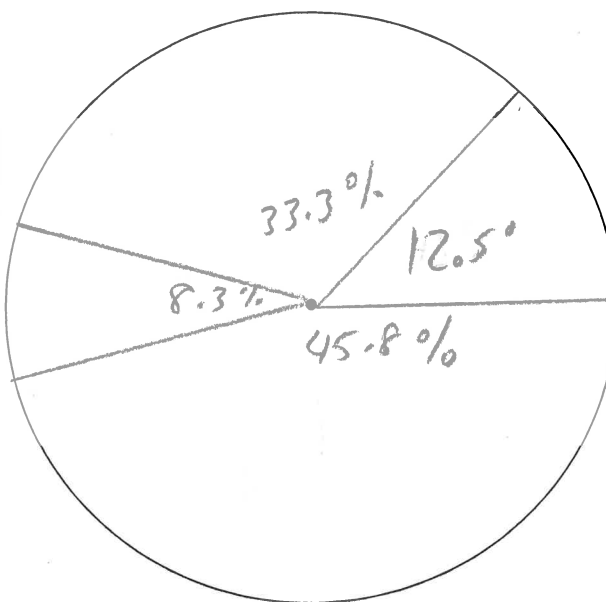
Steps to Draw an Angle



remember
 < greater than 90 look like this ↗
 < less than 90 look like this ↘

How your Day was Divided

45.8% at home (11hr)
 33.3% at work (8hr)
 12.5% in car (3hr)
 8.3% at gym (2hr)



Legend
 stating what
 sectors
 represent

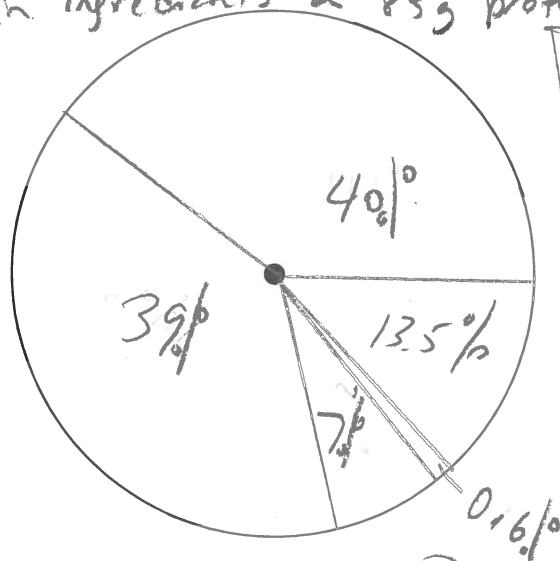
Try it! Create a circle graph to represent the following data. Liza is a fitness coach. She read the main ingredients in a new 85.0g protein bar.

a) Complete the chart.

"Chocolate Brownie" Protein Bar			
Ingredient	Mass (g)	% of total mass	Angle measure
protein	34.0	$\frac{34}{85} \times 100 = 40\%$	$0.4 \times 360 = 144^\circ$
total carbohydrates	33.0	$\frac{33}{85} \times 100 = 39\%$	$0.39 \times 360 = 140^\circ$
total fat	6.0	$\frac{6}{85} \times 100 = 7\%$	$0.07 \times 360 = 25^\circ$
sodium and potassium	0.5	$\frac{0.5}{85} \times 100 = 0.6\%$	$0.006 \times 360 = 2^\circ$
other	11.5	$\frac{11.5}{85} \times 100 = 13.5\%$	$0.135 \times 360 = 49^\circ$
total	85	100.1	360°

b) Draw a circle chart of the above data. Include a legend.

Main Ingredients in 85g protein bar



40% (34g) protein
39% (33g) carbs
7% (6g) fat
0.6% (0.5g) sodium/
potassium
13.5% (11.5g) other

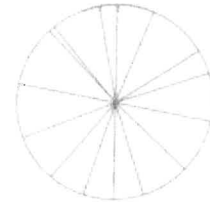
c) The mass of protein is more/less than that of carbohydrates.

d) Protein has almost 6 times more mass than total fat.

e) What other type of graph could you use for this data? bar

Some Advantages and Disadvantages of Circle Graphs

<u>advantages</u>	<u>disadvantages</u>
<ul style="list-style-type: none"> • Present data in a visual way that makes it easy to compare different categories of data (<i>see example p,31 - easy to see how paycheque spent</i>) • Visually simple: Great way to summarize results in a way that is easily understood: largest is majority; smallest is minority • Eye catching and easy to understand 	<ul style="list-style-type: none"> • Not good for showing relationships between two or more sets of data • Normally show percentages and not actual data values and therefore can be easily manipulated to give false impressions (<i>example - if you asked 50 girls and 200 boys a question and 25 girls and 100 boys said "yes", it would look like an equal amount of people said "yes" because each category would be 50%</i>) • Not very effective if there are too many sectors. (<i>ex - can you easily tell which section is the largest and which is the smallest? </i>) • If there are too many sections, you can use a category of "other" for the small sections, but then you wouldn't know what data is in the "other" section

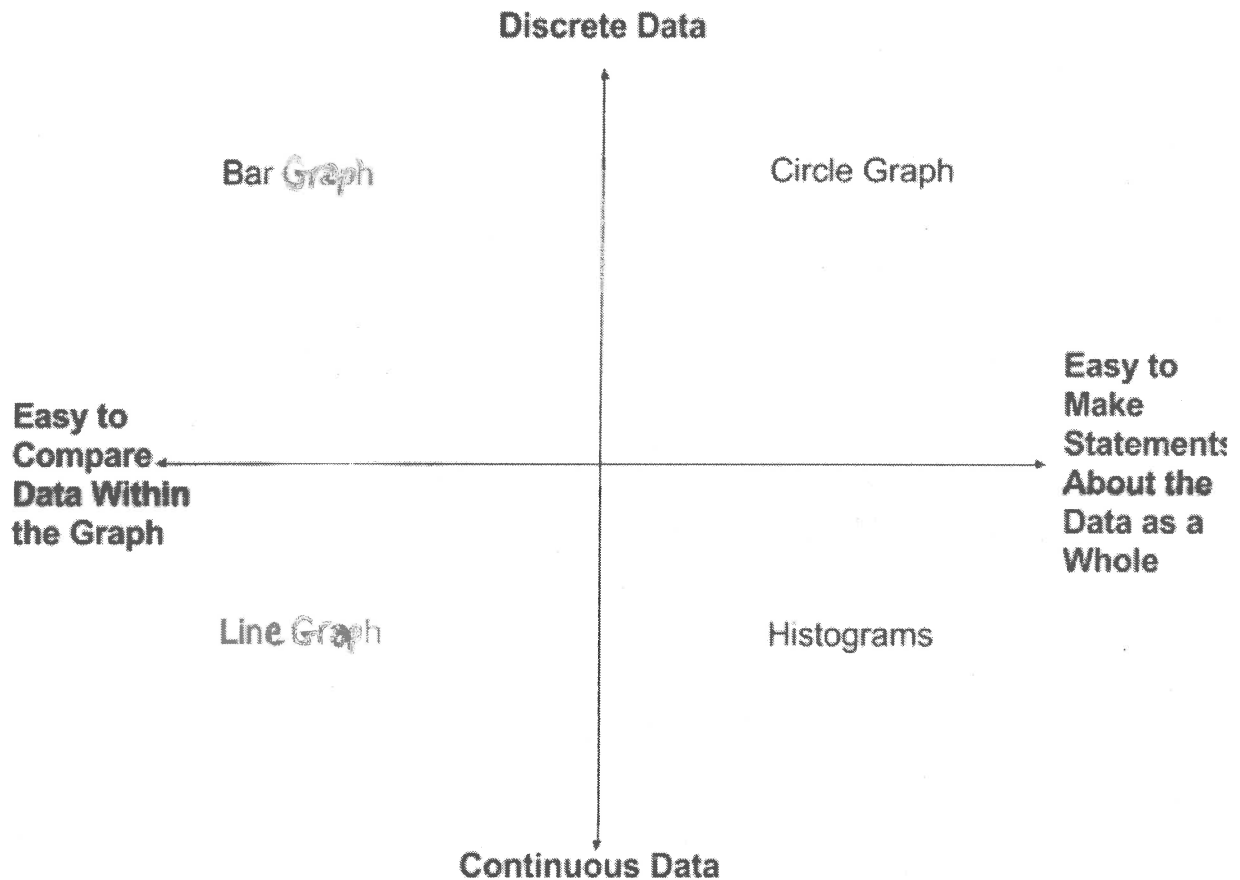


Comparing Types of Graphs - Use this figure below to help you to choose which type of graph to use for your data.

The Compass Rose of Graphical Interpretations

Each segment of the figure below contains one of the graph types we have talked about. The points of this compass are labelled as follows:

- **Discrete Data:** the graph contains discrete data
- **Continuous Data:** the graph contains continuous data
- **Easy to Compare Data Within the Graph:** the design of the graph makes comparison of information within the graph simple
- **Easy to Make Statements About the Data as a Whole:** the design of the graph allows for statements about the whole data set simple to see



How to Choose the Graph to Use

Graphs help us examine trends and make comparisons by visually displaying data. Before we can graph a given set of data from a table, we must first determine which type of graph is appropriate for summarizing that data. We have seen several types of graphs, each with its own purpose, and its own strengths and limitations. Which one we choose depends on the type of data given, and what we are asked to convey to the reader.

Sneakers Sold This Month	
Brand	Number Sold
Adidas	25
New Balance	18
Nike	32
Reebok	15
Other	10

Example 1: The table shows the number of sneakers sold by brand for this month. Construct a graph which best demonstrates the sales of each brand.

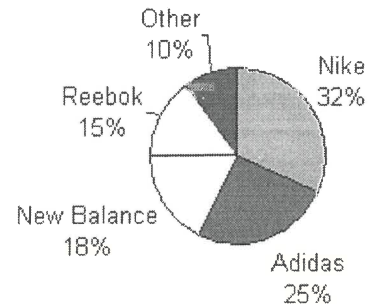
Analysis:

The numerical data in this table is not changing over time. So a **line graph would not be appropriate** for summarizing the given data.

Let's draw a circle graph and a bar graph, and then compare them to see which one makes sense for this data. Before we can draw a circle graph, we need to do some calculations. We must also order the data from greatest to least so that the sectors of the circle graph are drawn from largest to smallest, in a clockwise direction.

Sneakers Sold This Month				
Brand	Number Sold	Percent	Decimal	Angle Measure
Nike	32	32	0.32	$0.32 \times 360^\circ = 115.2^\circ$
Adidas	25	25	0.25	$0.25 \times 360^\circ = 90^\circ$
New Balance	18	18	0.18	$0.18 \times 360^\circ = 64.8^\circ$
Reebok	15	15	0.15	$0.15 \times 360^\circ = 54^\circ$
Other	10	10	0.10	$0.10 \times 360^\circ = 36^\circ$
Total	100	100%	1.00 = 1	360°

Sneakers Sold This Month



Circle graphs are best used to compare the parts of a whole.

The circle graph above shows the entire amount sold. It also shows the percent of each brand's sales as part of that whole, but doesn't show the numbers of each sold.

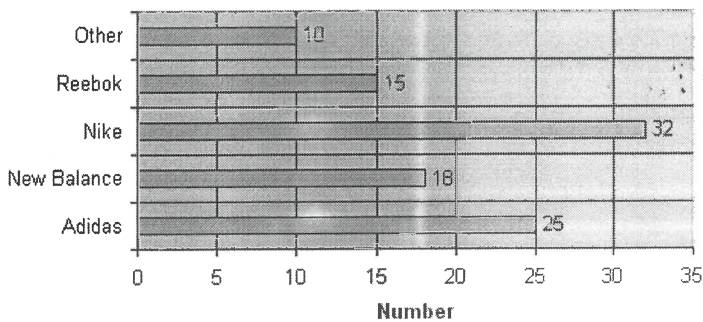
The circle graph uses the total of all items in the table.

Each sector of the circle graph is in the same proportion to the whole circle as the number of sales for that industry is to the entire amount of sales from the table.

To construct an accurate circle graph, you must first order the data in the table from greatest to least. You also need to find each part of the whole through several elaborate calculations and then use a protractor to draw each angle.

We were asked to show the individual sales of each brands; not the percent of the whole. If we were asked to show that the Nike brand dominates the sneaker industry, then the circle graph would be a better choice for summarizing this data.

Sneakers Sold This Month



Bar graphs are used to **compare facts**.

The bar graph stresses the individual sales of each brand as compared to the others.

The bar graph does not use the total of all items in the table.

The bar graph simply gives a visual listing of the information in the table.

The number of sneakers sold for each item in the table matches the value of each bar in the bar graph. This makes the bar graph a more direct and accurate way of representing the data in the table.

We were asked to construct a graph which best demonstrates the sales of each brand.

Solution: Each graph above has its own strengths and limitations. However, the **bar graph is the best choice for summarizing this data based on what we were asked to convey to the reader.**

Example 2: The table below shows the humidity level, recorded in Small Town, NY for seven days. Construct a graph which best demonstrates the humidity level for each day.

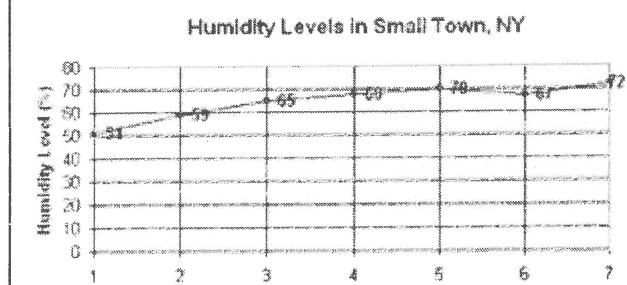
Humidity levels in Small Town, NY

day	Humidity level (%)
1	51
2	59
3	65
4	68
5	70
6	67
7	72

Analysis:

The humidity level is given as a percent. At first glance, it might lead one to think that a circle graph should be used to summarize this data. However, the data in the table **does not** include any parts in relation to a whole. Thus, a circle graph is not the right choice. The **data in this table is changing over time**. So a **line graph** would be the appropriate choice for summarizing the given data.

Solution:



Example 3: The table shows the composition of Earth's atmosphere. Construct a graph which best represents the composition of the Earth's Atmosphere.

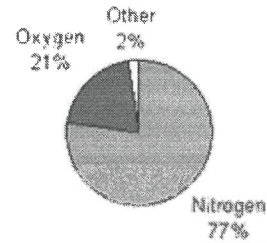
Composition of Earth's Atmosphere	
Gas	percent
nitrogen	77
oxygen	21
other	2

Analysis:

The word "composition" indicates that we are looking at the parts of a whole. The Earth's Atmosphere is the **whole** (100%) and each gas is a **part of that whole**. Accordingly, a circle graph is the best choice for summarizing this data.

Solution:

Composition of Earth's Atmosphere



Try It!

For each of the following, should you use a circle graph, bar graph, or line graph? Why?

1. The ages of 7 trumpet players in a band are 13, 12, 11, 12, 11, 10, and 12. What type of graph would be appropriate for comparing the ages of these trumpet players? bar (or histogram)

*age of trumpet players discrete, separate facts
each bar represents an age of a player
single number*

2. The federal hourly minimum wage was recorded each year from 1990 to 2007. What type of graph would best show the changes in minimum wage during this time period? line

date changing over time

3. When asked if "antidisestablishmentarianism" has 28 letters, 50 people said "yes", 35 people said "no", and 15 people said "I don't know". What type of graph would best compare these responses to each other and with the total? circle

all people - total 50/100 35/100 15/100

4. The growth of 7 different plants was recorded in centimetres. What type of graph would be best for comparing the growth of each plant? bar

discrete, separate parts

5. In a city, the rainfall was recorded in inches each month for 12 months. What type of graph would best display the change in rainfall? line

date changing over time

Comparing Types of Graphs:

7, 15, 24, 31
p. 13, 22, 30, 35, 37-39

There are four types of graphs, but when do you use each type???

Type of Graph	When to use it...
Bar Graph	<ul style="list-style-type: none"> - discrete data that is clearly separate from each other - compare data, facts, well-defined separate categories within graph <p>example: students' favorite fruit</p>
Histogram	<ul style="list-style-type: none"> - continuous data - make statements about data as a whole - range of data all related to one topic - summarize data with number grouped in ranges <p>example: tree heights in orchard - # of trees in each category of height</p>
Line Graph	<ul style="list-style-type: none"> - continuous data - compare data within graph - data changing over time - 2 groups of data with a relationship where 1 affects the other - see trends in data - can interpolate, extrapolate <p>example - average temp each month in 2016</p>
Circle Graph	<ul style="list-style-type: none"> - discrete data - make statements about data as a whole - categories of data that are parts of a whole (1% of whole) <p>example: % of total that liked each kind of fruit</p>

Lesson #6: Interpreting Graphs

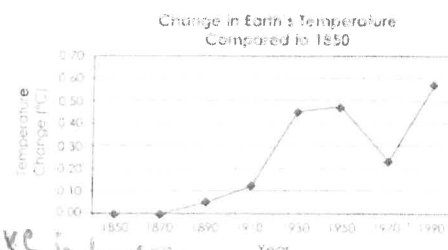
We have seen how to draw four different types of graphs. These four types - bar, histogram, line, and circle - are frequently used by media, advertisers, and others to provide statistical information in a way that can be quickly and easily understood.

However graphs are sometimes used to **influence other people by** modifying the way a graph is drawn. Even sometimes an "honest" graph can present misleading information.

Multiple Interpretations of One Graph

- Your interpretation of a graph may be determined by your **personal beliefs** or by the **factual data** (numbers) presented by the graph.

The following graph describes the average temperature changes on the earth from 1850 CE to 1980. Based on the information in the graph, we could draw many different conclusions about Earth's temperature. For example:



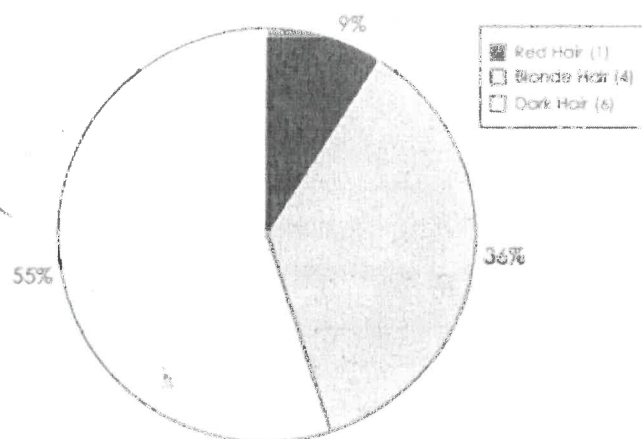
- personal beliefs*
1. Whatever we did between 1950 and 1970 was effective in lowering Earth's temperature.
 2. As it did between 1930 and 1950, Earth's temperature may plateau (stay approximately the same).
- factual data*
3. Earth's temperature is increasing rapidly.
 4. We do not need to be concerned about Earth's temperature because it is not increasing by that much.

The first two statements might be personal beliefs based on whether someone is concerned or not about an increase in the Earth's temperature. The second two statements might be based on the factual data. One person might be concerned about the sharp vertical rise. Another might not be concerned about a temperature increase of 0.6°C in 100 years.

Try it! State two different ways someone could interpret the graph. Include one possible reason for each interpretation.

1. most people have dark hair
50% of population - largest part of circle
2. blond and dark hair almost equally common
4 blonde 6 dark - note a large difference

Hair Colour of Kids at Sunshine Day Care

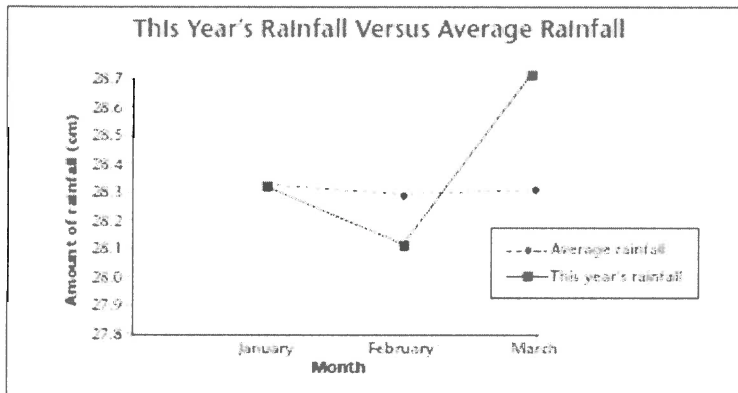


Misleading Data:

Graphs can be used to display your data at a glance. However, graphs can distort your results if you are not careful. The picture that results may not be objective, or without bias or distortion. Look at the first graph.

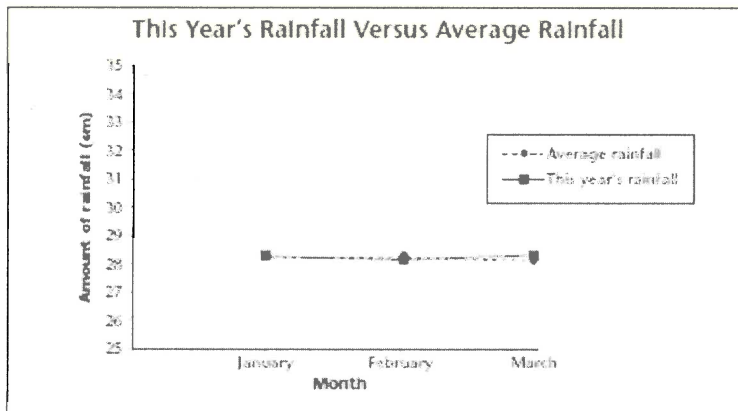
How Much Rain Really Fell?

In the graph below, it appears as though March had drastically more rainfall compared with an average month. But did that really happen?



Wait! March's rainfall was only 0.4 cm above average. On the graph, that looks like a large increase. On the ground, a 0.4 cm increase is not that much. This graph is biased because it exaggerates the difference between the two lines. Because the interval between 27.8 cm to 28.7 cm on the y-axis is so small, the difference in rainfall seems very large and noticeable.

If you increase the interval between numbers on the y-axis, the scale becomes larger. That makes the difference between the two lines smaller, as shown below.



1. What is the range of values on the y-axis in the second graph?

25-35

2. Which graph is a more accurate picture of the data? Explain.

second.

0.4cm is a very small increase
2nd graph shows that.

See p. 30

AT + T

MCI

graphs

+ p. 13

how to mislead
with scale or
size of
category

Smaller

scale
exaggerated

the

small

differences

and

biased

the

graph

A Matter of Scale

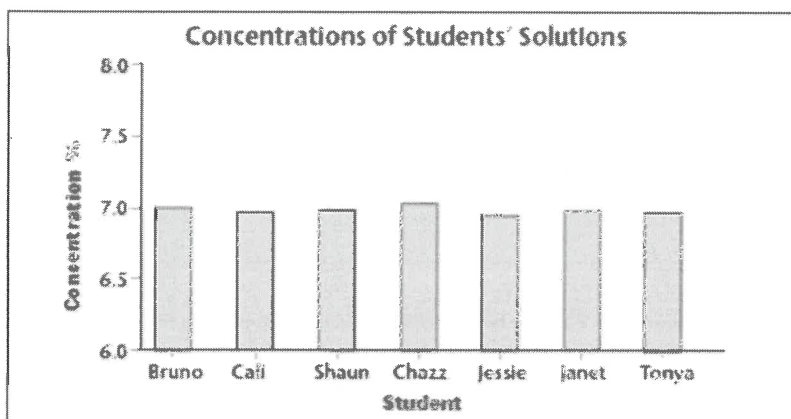
Here is another example of how the choice of the scale can alter a graph.

In an experiment, seven students tried to mix a solution of salt water so that its concentration would be exactly 7.00%. When the teacher tested the concentration of their solutions, he got the following results:

Concentrations of Students' Solutions

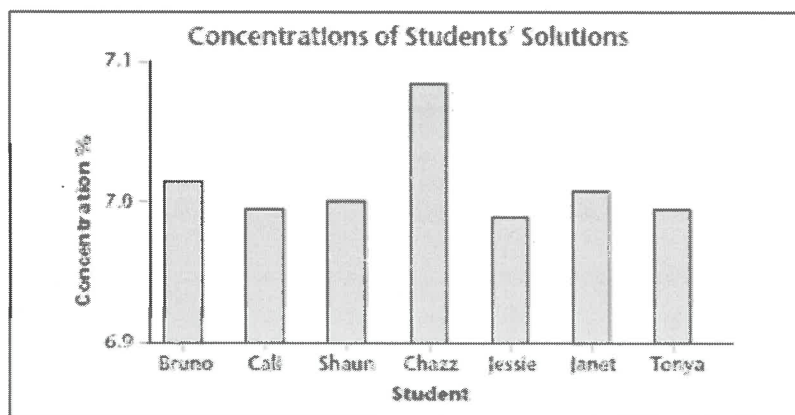
Name	Bruno	Cali	Shaun	Chazz	Jessie	Janet	Tonya
Concentration	7.02%	6.99%	7.00%	7.08%	6.97%	7.01%	6.99%

The teacher created the following graph to show the students' results:



Does this graph give you a clear picture of how the concentrations varied? Not really. The bars look so much alike that it's hard to tell the differences between them.

Suppose the teacher decreased the scale of the y-axis. The graph would then look like the one below. The variation in the students' results looks much greater, even though it hasn't changed. This graph makes it easier to see the small differences.



*Smaller
Scale
Easy to
see
small
differences
if needed*

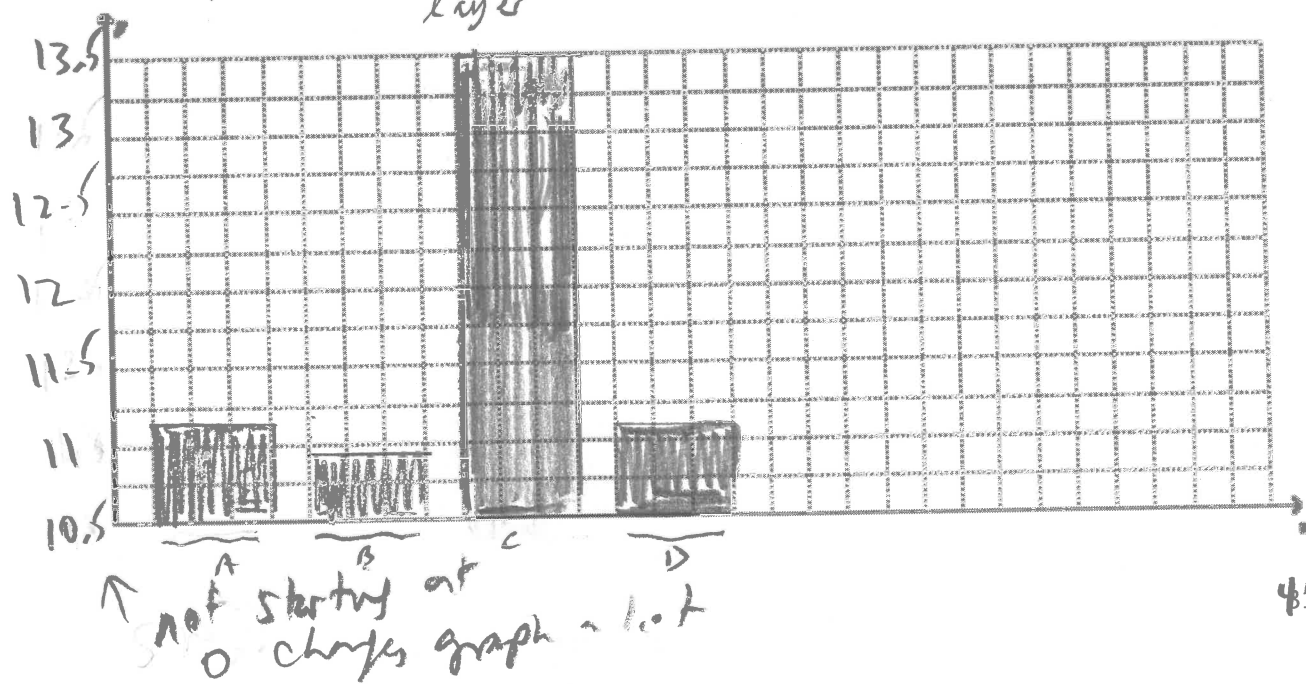
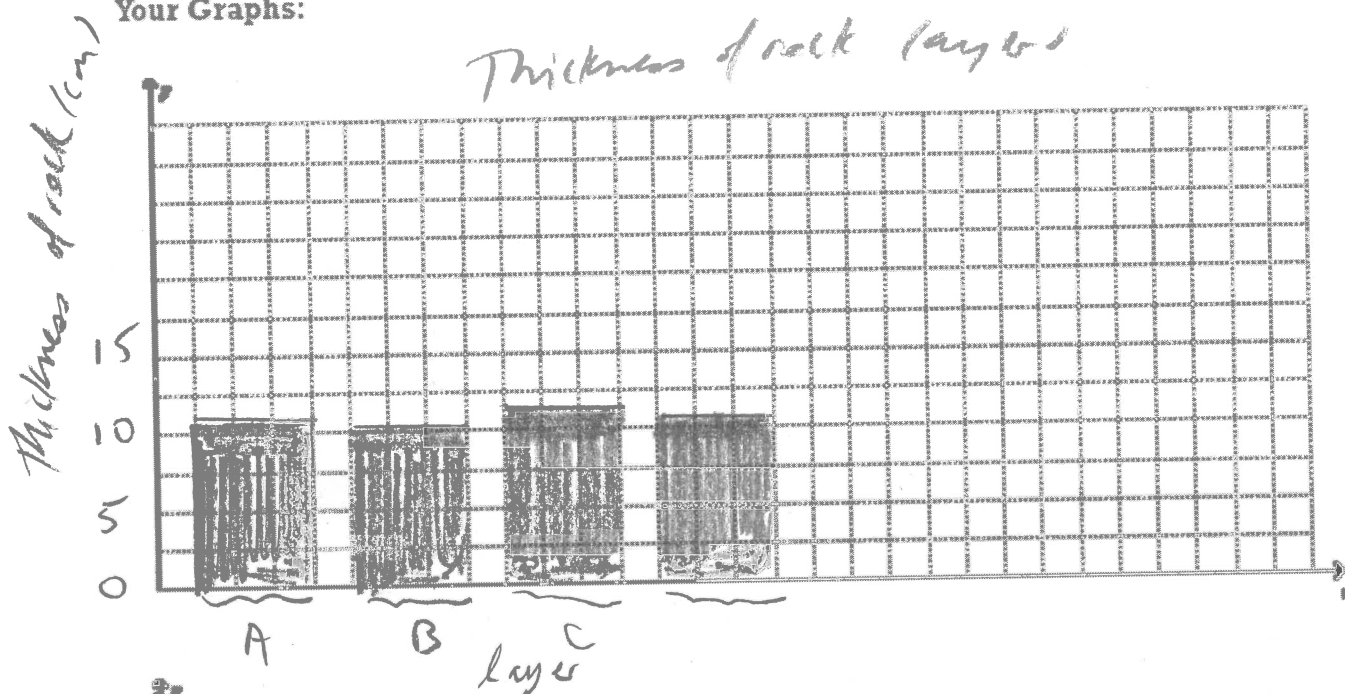
Graphs with an Attitude

The data in the chart below were recorded by a student measuring the thickness of four rock layers.

Thickness of Rock Layers

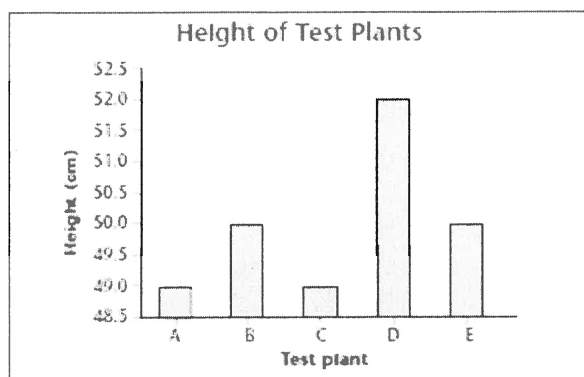
Layer	A	B	C	D
Thickness	11.2	10.8	13.5	11.1

Using the above data, create two graphs in the space below. First show how similar the measurements are. (Hint: Make the scale of the y-axis larger. This makes the difference between the measurements seem smaller.) In your second graph, emphasize the fact that layer C was slightly thicker than the other layers.

Your Graphs:

Identifying Bias on Your Own

Graph 1

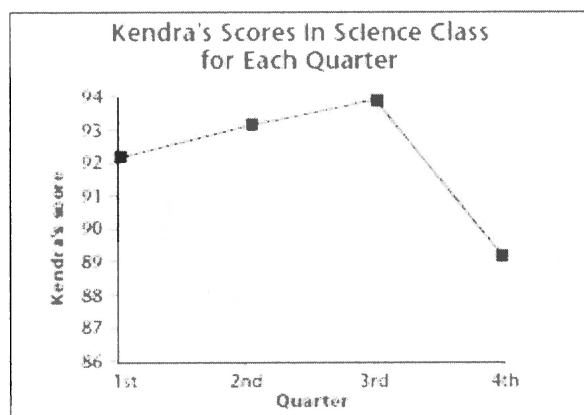


A 49
B 50
C 49
D 52
E 50

1. This graph shows that test plant D grew much taller than the other plants. How is this information misleading?

D is only 2-3 cm taller than others. That's a little taller but not much if it's a big plant.

Graph 2

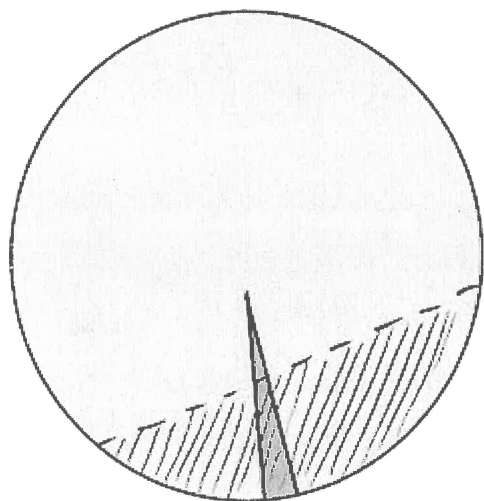


92 → 93 → 94 → 89.

2. This graph shows that Kendra received a much lower grade in science class during the fourth quarter. Do you think what appears to be such a large drop in her grades should worry Kendra? Explain your reasoning.

It is a 5% drop but it's still pretty good.

A graph on math jokes



□ Math jokes that are funny

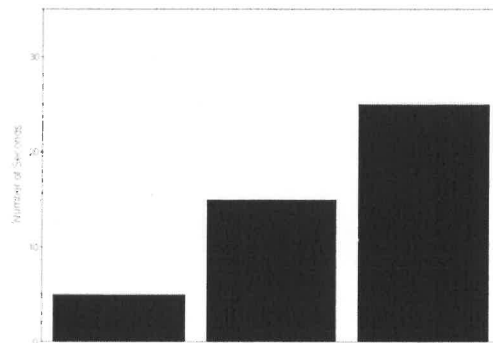
□ Math jokes that are not funny

/// math jokes that you could actually understand



a spider

How Long It Took Me To Draw and Color In Each Of These Bars



©2011 Ben Crutcher's Student Notes



kappit.com